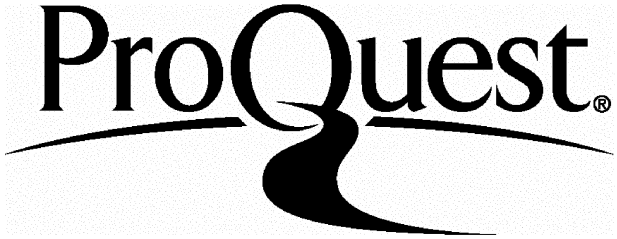
Identification and quantification of metabolising enzymes in human tissue: a proteomics approach

Cathrin Seibert

A thesis submitted for the partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Pharmaceutical and Biological Chemistry
School of Pharmacy
University of London

March 2008



Plagiarism statement

This thesis describes research conducted in the School of Pharmacy, University of London between March 2004 and March 2008 under the supervision of Professor William J. Griffith. I certify that the research described is original and that any parts of the work that have been conducted by collaboration are clearly indicated. I also certify that I have written all the text herein and have clearly indicated by suitable citation any part of this dissertation that has already appeared in publication.

Cathi Seibert

Signature

14. August 2008

Date

Abstract

Metabolism is a complex process, taking place in every cell of the organism, comprising a variety of different chemical reactions. Knowledge about enzymes involved in metabolism of endogenous and exogenous substances is crucial for the understanding of detoxifying and toxifying processes in an organism. These in turn can play a role in the generation of a diseased state or contribute to healing. From this knowledge one can also predict e.g. the capability of a diseased tissue to metabolise an administered drug. This information supports the development of drug treatments, as different expression profiles of metabolising enzymes can be expected in different tissue types. The most prominent group of metabolising enzymes are presented by the Cytochrome P450 family; in humans there are 57 functioning genes known to date.

Analysis of metabolising enzymes was performed using mass spectrometry. Thereby detection of enzymes is possible directly at the protein level. Microsomal samples were submitted to one-dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis, followed by in-gel tryptic digestion. Extracted peptides were then analysed using nano-liquid chromatography (nano-LC) coupled to nano-electrospray tandem mass spectrometry (ES-MS/MS) and protein database searches were used for protein identification. Emphasis of the initial part of this work was on optimisation of the mass spectrometry method in order to increase the number of identified proteins in human lung microsomes. This was achieved by dividing the mass spectrometry scan range into smaller ranges. Hereby a 1.5-times higher number of proteins was identified

In a quantitative approach, seven human liver tissue samples were analysed and the content of Cytochrome P450 enzymes (CYP2E1 and CYP1A2) determined. This was accomplished using a stable isotope labelled internal standard peptide. The native and synthetic peptides show the same physical properties for analysis by LC-MS/MS, such as retention time in the chromatogram and are only differing in their molecular weight. This method was validated to prove reproducibility and linearity of the quantification technique. Quantification of CYP2E1 resulted in 88–200 pmol per mg of microsomal protein; for CYP1A2 165–240 pmol/mg microsomal protein were detected.

Acknowledgements

I would like to thank all people who have helped and inspired me during my doctoral study.

I express my sincere appreciation to my first supervisor, Professor William J. Griffiths, for his advice, guidance and support during my studies and for giving me the opportunity to present my work at scientific conferences. I would also like to thank my second supervisor Professor Laurence H. Patterson for his help and advice.

Thanks also go to Dr. Yuqin Wang for training on the QTOF and her help in proteomics questions.

The financial support of the School of Pharmacy is gratefully acknowledged.

I am deeply grateful to Dr Luciano Saso and Professor Andreas Kortenkamp for their encouragement to pursue a PhD.

A big thank-you to all my colleagues in the mass spectrometry group: Ahmed Al-Gazzar, Richard Betts, Kersti Karu, Dr. Cathy Lane, Mathangi Narayanaswamy, Dr. Sonia Nisar, Emmanuel Samuel and Amar Sharif.

Thanks also to Dr. Mire Zloh for his help.

Many thanks to Dr. Andy Wilderspin for using his lab and giving advice on gel problems.

I also thank Derek Marley for his much appreciated company during undergraduate teaching and for always having an open ear.

I am indebted to my friends support here in England as well as in Germany and the rest of the world but it is difficult to mention them all.

Special thanks are due to Sibylle Heidelberger for her friendship and help in the lab and enthusiasm at work.

I am very grateful to Anique Raether for her companionship over the last ten years as a colleague, flatmate and friend.

I owe most sincere gratitude to Dr Gunter Kuhnle for all his help and the incredible amount of understanding and patience he had with me especially in the last six months.

I would like to express my deep and sincere gratitude to my whole family. I would not be here without the love and support of my grandparents and my parents, Ursula and Karl-Gerd Seibert, and their unconditional belief in me. — Opa Kurt, I could not have done this without you.

Table of contents

List of figures

List of tables

Abbreviations

1D	One dimensional
2D	Two dimensional
ANOVA	Analysis of variance
APS	ammonium-persulphate
Bis-Tris	Bis-(2-hydroxyethyl)-imino-tris-(hydroxymethyl) methane-HCl
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
cps	counts per second
CYP	Cytochrome P450
CYPxyz	Specific cytochrome P450; x denotes family, y subfamily and z polypeptide
CV	Coefficient of variation (standard deviation/average)
Da	Dalton
DDA	data dependent analysis
DTT	Dithiotreitol
E. coli	Escherichia coli
ESI	Electrospray ionisation
GluFib	human [Glu1]-Fibrinopeptide B
GO	gene ontology
HEPES	4-(2-hydroxyethyl)-1-piper-azine-ethanesulfonic acid
HPLC	High-performance liquid chromatography
HSE	Health and Safety Executive (UK)
isp	internal standard peptide
LC	liquid chromatography
LC-MS	liquid chromatography coupled with mass spectrometry
LC-MS/MS	liquid chromatography coupled with tandem mass spectrometry
m/z	mass-to-charge ratio
MALDI	Matrix-assisted laser desorption ionisation

MBP	maltose-binding protein
MOPS	3-(N-mor-pholino) propane sulfonic acid
MRM	Multiple reaction monitoring
MS	mass spectrometry
MS/MS	tandem mass spectrometry
mRNA	messenger RNA
MW	Molecular weight
nr	non-redundant
PAGE	polyacryl amide gel electrophoresis
PCR	Polymerase chain reaction
Protein score	-10log10(P); (P is the probability that the observed match between the experimental data and the sequences in the database is a random event)
Q-TOF	Quadrupole-time-of-flight mass analyser
QUAD	Quadrupol
RNA	Ribonucleic acid
RT-PCR	Real-time PCR
RP-LS	Reversed-phase liquid chromatography
SDS	sodium dodecyl (lauryl) sulphate
SEM	Standard error of the means
Stdev	Standard deviation
TEAB	triethylammoniumbicarbonate
TEMED	tetramethylethylenediamin
TRAP	Quadrupole on trap
Tris	Tris-(hydroxy-methyl)-aminomethane
TOF	Time-of-flight
UKHTB	UK Human tissue bank

Amino acids

Amino acid	One-letter code	Three-letter code
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cystein	C	Cys
Glutamic acid	E	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val
Aspragine or aspartic acid	B	Asx
Glutamine or glutamic acid	Z	Glx

To my parents and

grandparents

Chapter 1 Introduction

The regulation of physiological processes is the subject of extensive studies. An important part of this research is dedicated to the investigation of the gene expression profile at the protein level, the "proteome". This term was first used by Mark Wilkins in 1994, who was looking for a word to describe the increasing number of proteins being identified at that time [1]. Proteomics, the study of the proteome, has emerged as a rapidly growing area and is used in various fields of research. [2, 3]. To understand the development of diseases and their origin it is important to be aware of the expression patterns of proteins in healthy and diseased tissue. Differences in the protein-pattern of a cell or tissue-type can reveal significant information not only for the understanding of a disease, but also to find a starting point for possible treatments. Enzymes involved in metabolism represent a group of proteins, which are of interest when investigating diseases and developing methods for treatment. Endogenous as well as xenobiotic compounds are metabolised by these enzymes, resulting in either detoxified or even the generation of more toxic forms [4], which can in turn have influence on the generation of the disease state or contribute to healing. Knowledge about these enzymes is therefore crucial and proteomics provides powerful tools to pursue the identification of proteins.

1.1 Proteomics

Proteomics is the analysis of all proteins of a cell, or tissue type at a particular moment in time. It is an attempt to study biological processes beyond the level of genes by investigating the structure and function of proteins and their involvement in various pathways [5]. Proteomics finds its application in various fields of current research [6], such as disease diagnostics [7], drug discovery [8] or specialised areas such as tissue imaging [9]. A significant aim in proteomics is to be eventually able to identify the proteome of any organ or tissue type in any organism [10]. The human proteome organisation (HUPO, www.hupo.org) for example, tries to arrange a combination of proteomics studies worldwide in order to gather information about various human organ, tissue and fluid

proteomes [11]. The proteome of the human brain [12], liver [13] and plasma [14] have been tackled so far. The cooperation of different individual research groups in these projects is necessary to cover the wide-range of expressed proteins, which have to be identified and systematically ordered. There is no common agreement on the exact number of genes in the human genome, but numbers of around 25,000 genes (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml) have been suggested. One gene can code for more than one protein because of (alternative) mRNA splicing, as well as editing and changing of the reading frame in the genome. Investigation of gene expression was traditionally performed on the mRNA level, performing DNA microarrays or real time PCR [15, 16]. However, messenger RNA abundance does not necessarily correlate with protein abundances in a cell [17] because translation and transcription are two distinct processes which are regulated separately. Furthermore the proteome, i.e. the **prote**in complement expressed by a gen**ome**, differs from cell to cell. It underlies a dynamic change for example in the types of expressed proteins, their abundance, sub cellular location, modification or function [18]. Just mentioning the over 300 physiological possible posttranslational modifications [19] implies a multitude of different products, as many of the twenty proteinogenic amino acids can be modified in various ways, including e.g. phosphorylation, glycosylation, acetylation and sulphatation. There are estimates of about '10 to 20 million distinct polypeptides' potentially present in a single tissue type [1]. Thus, the knowledge of the proteome is of importance, as the information of gene sequences is not sufficient to determine the protein-character of a cell. Nevertheless genetic information is needed to predict and confirm the proteomic research outcome.

1.2 Tools of proteomics

1.2.1 Gel-electrophoresis

Gel-Electrophoresis is a commonly used technique to separate proteins according to their properties of molecular weight or isoelectric point. This technique relies on the migration of charged protein molecules in a matrix of polyacrylamide upon application of an electric field [20, 21].

Both one or two dimensional gel electrophoresis are possible. In this work proteins were separated by a one-dimensional SDS PAGE system (sodium dodecyl sulphate polyacrylamide gel-electrophoresis). The principle of this system is to denature proteins first by heating to 100°C e.g. with DTT (dithiotreitol) and an excess of SDS. DTT is used to break existing disulphide bonds in proteins. The hereby formed free SH-groups are alkylated with iodoacetamide to prevent reformation of disulphide bonds [22]. SDS is a negatively charged, denaturing detergent, which can bind to proteins forming SDS-polypeptide complexes in a constant weight ratio. The negative charges of the SDS detergent overweigh the intrinsic charges of the polypeptide chains. Thus, these chains have the same charge densities and differ only in their molecular weight, which leads to a migration in the pores of a polyacrylamide gel depending only on their weight. The pores of these gels are created by free-radical triggered polymerisation of monomeric acryl amide and bis-acrylamide (N,N'-methylarlene-bisacrylamide) [23]. Depending on the ratio of both, the resulting gel has different pore sizes, which influence the migration of molecules. Small pores hinder large molecules in migration, so the bigger the pores the better the conditions for high-molecular weight molecules to migrate in the gel. Hence the percentage of a gel is chosen according to the size of proteins to be analysed. As a discontinuous gel system was applied in our work, the samples were loaded on a stacking gel, which is polymerised on top of the resolving gel. The two gels differ in pore size, with larger pores in the stacking gel, and have a different pH (pH 6.8 in the stacking gel, pH 8.8 in the resolving gel). In the stacking gel, the proteins are focused in a narrow band between the fast moving chloride ions and the slow moving glycine which is neutral at pH 6.8. On entering the resolving gel, glycine becomes protonated and migrates faster than the proteins, whereas the proteins become separated according to their size [24, 25].

In two-dimensional gel electrophoresis, proteins are separated first according to their isoelectrical point, at which point the sum of the positive and negative charged amino acids is zero, so the protein shows no mobility in the electrical field. In the second dimension, proteins are separated by SDS-PAGE according to their molecular weight [26, 27]. For this work one-dimensional gel electrophoresis was chosen, as particular interest was in CYP proteins, which represent a membrane associated enzyme family of proteins. With two-dimensional gel-electrophoresis it has been difficult to detect

membrane proteins, which might be due to their hydrophobicity and basic pK [28]. Furthermore attempts to separate CYP proteins on two-dimensional gels were not successful [29]. After electrophoresis, proteins can be detected with silver stain, fluorescent dyes or, as used in this study, Coomassie brilliant blue stain [24].

1.2.2 Mass spectrometry

Mass spectrometry (MS) is an analytical technique used in a wide range of fields. For the study of peptides and proteins it has developed into a very powerful tool. Eugen Goldstein made an early discovery for mass spectrometry in 1886. He discovered the "canal rays", as he found that while cathodes emit cathode rays, another sort of rays follow the opposite direction [30]. Ten years later Sir John J. Thomson established that these canal rays consist of positively charged material and he built the first mass spectrometer, using parabolic deflections of 'positive rays' with photographic plates [31]. Francis Aston introduced the first mass spectrometer, which could focus the beam of ions with the same *mass-to-charge-ratio* (*m/z*) in one spot on a photo plate-detector in 1919 [32, 33]. Although initially developed for physico-chemical research, such as isotope separations, mass spectrometry was introduced into the chemical industry in the 1930s for product analysis, including petrochemistry. After it had been used for many years in these areas and mainly for the identification of small, volatile compounds, at the end of the 1950s first attempts were undertaken to investigate primary structures of proteins by mass spectrometry [34]. This was enabled by derivatising peptides into more volatile compounds, which then were analysed with gas-chromatography mass spectrometry (GC-MS). For the following 20 years protein identification via mass spectrometry was combined with the results derived from Edman degradation. In 1981, a soft ionisation technique, fast atom bombardment (FAB) ionisation, was developed by Barber and Bordoli [35] who achieved the ionisation of intact peptides without derivatisation. Seven years later Hillenkamp [36] developed the matrix-assisted laser desorption/ionisation technique and Fenn [37] in 1989 'invented' electrospray ionisation. These soft ionisation techniques paved the way for further more facile protein analysis.

Mass spectrometry in general is based on the ionisation of molecules in the ion source, separation according to their *mass to charge* ratio in the analyser, followed by their detection in proportion to their abundance (detector) (Figure 1.1) [38].

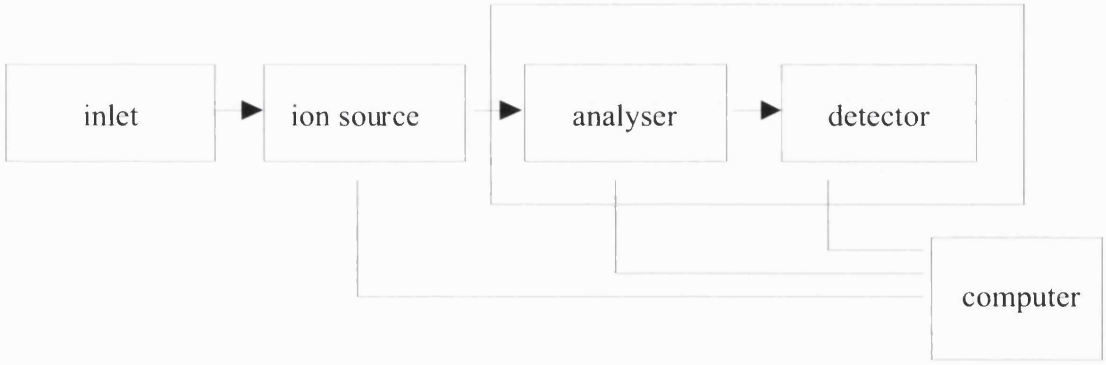


Figure 1.1: Schematic assembly of a mass spectrometer. Samples are introduced via the inlet device, ionised in the ion source and analysed in the mass analyser. The whole process as well as the detection of signals is controlled via computer equipped with suitable software.

To obtain additional information on structure or sequence, it is possible to fragment ions in the mass spectrometer. The fragments obtained can then be used to characterise the ion further. To fragment ions, a technique called Collision Induced Dissociation (CID) [39] is used. Sample ions travel through a chamber filled with a collision gas (often helium or nitrogen). When sample ions and gas molecules collide with sufficient energy, the sample ions start to dissociate and form so called product ions. Using appropriate collision energies, it is possible to dissociate peptides or small proteins at their peptide bond, thus providing sequence information. To select ions for collision induced dissociation, a mass filter (often a quadrupole, see below) is used; the product ions can then be analysed either with a second quadrupole, a quadrupole ion trap or TOF analyser. The combination of several mass analysers with a collision cell is commonly referred to as tandem MS.

1.2.2.1 Ionisation

In order to analyse a molecule with mass spectrometry, it has to be ionised. There are different ways to ionise a molecule. In proteomics, mostly ESI (electrospray ionisation) and MALDI (matrix assisted laser-desorption ionisation) are used. These methods are regarded as soft ionisation methods, because the molecule fragmentation during ionisation is minimised. In MALDI, samples are ionised out of a solid mixture with a

matrix. In this work ESI was used as the ionisation method and is therefore described in more detail in this section. The achievements by Fenn *et al.* [37, 40] have already been mentioned above. Already twenty years before them, Dole *et al.* [41] used an electrospray source to nebulise and ionise a solution of polstyrene and reported the production of macro ions with the electrospray technique.

The advantage of a liquid-state ion source is the possibility of coupling it with liquid chromatography (LC) separation methods. The molecules are not actually ionised in the ion source but the ions are preformed in solution (due to pH and electrochemistry in the spray capillary) and extracted into the gas phase by the electrospray process. For example, proteins or peptides are usually positively charged in an acidic medium. The sample passes through a capillary to which a strong electric field is applied. This field, often aided by a stream of gas, nebulises the sample and transfers the ions from the liquid phase into the gas phase before they enter the vacuum system of the mass analyser [42]. Despite extensive research, the process of ionisation in electrospray is still not totally understood [43]. A schematic illustration of the process of ionisation is given in Figure 1.2. To the end of the capillary, a potential is applied (3–6 kV); in the case of peptide analysis usually a positive potential is used. This affects the sample solution, which is pumped through the capillary, and positive charged ions are drawn out of the tip to form a so-called 'Taylor cone'. Between the capillary and the counter electrode, the ions in solution experience an electrical field and are thus drawn towards the negative counter electrode. Thereby the Taylor cone is more and more drawn into a filament, where surface tension as well as electrostatic forces are present. As soon as the electrostatic forces exceed the surface tension, droplets are formed. On their way to the counter electrode, these droplets can be exposed to elevated temperatures either in a heated capillary or in the form of a heated counter stream of nitrogen gas. Thereby, the solvent in these droplets evaporates which makes them shrink. The charge density on the surface of the droplets increases until electrostatic (Coulomb) forces overcome the surface cohesion forces resulting in the disintegration of the droplets (Coulomb explosion). The point at which this happens is described as the 'Rayleigh limit' [44]. By this way continuously smaller and smaller droplets are produced. When the electrical field on their surface is high enough, ions are desorbed. The process of ion formation out of the droplets is currently being discussed and there are two theories about the final events

in electrospray [45]. There is the charge residue mechanism [41, 46], which assumes a continuous evaporation of solvent and formation of smaller droplets until only one unsolvated ion remains. The evaporation mechanisms [47] on the other hand, proposes evaporation of an ion from the surface of a droplet caused by the strong electric field at the droplet surface. Compounds with a high concentration at the surface are more likely to be desorbed. Large molecules (in general > 1000 Da) with several ionisable sites produce multiply charged ions, also because of the accumulation of charges in the droplets. The detection of these ions can be achieved even with analysers with limited *m/z* ranges as it is the mass to charge ratio (*m/z*) which is measured. The peaks usually obtained with ESI are $(M+zH)^{z+}$ in positive ion mode or $(M-zH)^{z-}$ in negative ion mode. Thus a mass analyser with a mass range of up to 2000 *m/z* can for example detect a triply charged insulin molecule (MW: 5807 Da, $[M+3H]^{3+}$; 1936.7 *m/z*).

The quality of the electrospray depends on several parameters: the spray voltage, the flow rate, the solvent, the source design and the composition of the sample. For higher flow rates, the spraying process is often facilitated by a nebuliser gas (pneumatically assisted electrospray). Electrospray is a concentration dependent technique and in general sensitivity improves with lower flow rates. The low amount of sample often available in proteomic experiments and the improved sensitivity make low flow rate ESI a method of choice in proteomics. In 1994, Wilm and Mann developed the nano-spray device [48] and they managed to produce droplets of a diameter of about 200 nm with flow rates less than 25 nL/min. In this way droplets of high surface to volume ratios are formed, which enabled the desorption of a higher number of analyte molecules [49]. As nano-electrospray sources are designed to opperate at flow rates less than 1 µl per minute, they are therefore used together with nano- and micro- liquid chromatography (1.2.4).

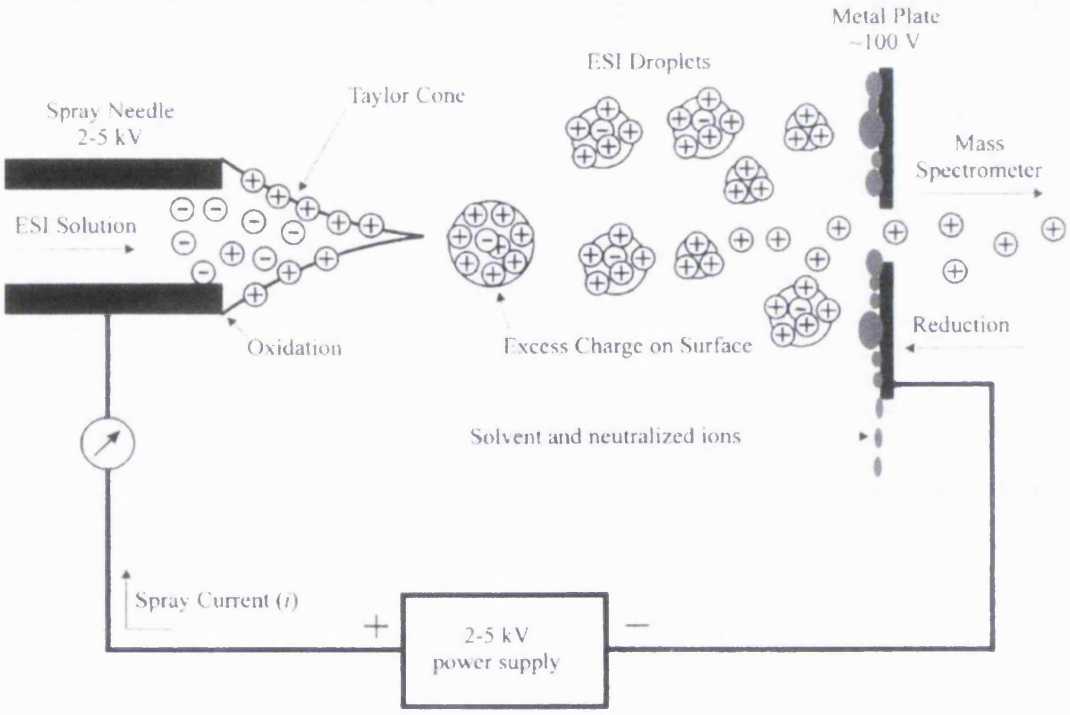


Figure 1.2: Schematic reproduction of the ionisation process in electrospray from Cech and Enke [50]. Solution of sample is introduced via a needle into the electrospray source. As the sample is charged and a voltage is applied on the outside of the needle, ions of opposite charge are drawn towards the needle surface: here negative charges in the solvent are drawn towards the positive voltage applied on the needle. A Taylor coned is formed between the needle and the counter electrode. Droplets are formed and move towards the negatively charged counter electrode whilst solvent evaporates. The moment surface tension on the droplets surpasses the electrostatic cohesion forces in the droplet (the Rayleigh limit), the droplets 'burst' (Coulomb explosion) and charged molecules are set free into the vacuum system of the mass spectrometer.

1.2.2.2 Mass Analysis and Detection

There are various types of analysers to separate the ions produced in the ion source. For this work the QTOF, a combination of quadrupole and time of flight mass analysers, as well as the quadrupole ion trap are of relevance. These instruments can be used in single

MS-mode, performing MS only, or MS/MS mode, performing tandem-mass spectrometry.

Quadrupole

The principle of the quadrupole mass filter is based on the work of Wolfgang Paul [51]. Simply, a quadrupole consists of four parallel rods with a hyperbolic cross section although a circular cross section can also be used. An electric field is applied to the rods, which leads to the oscillation of ions passing this electric field.

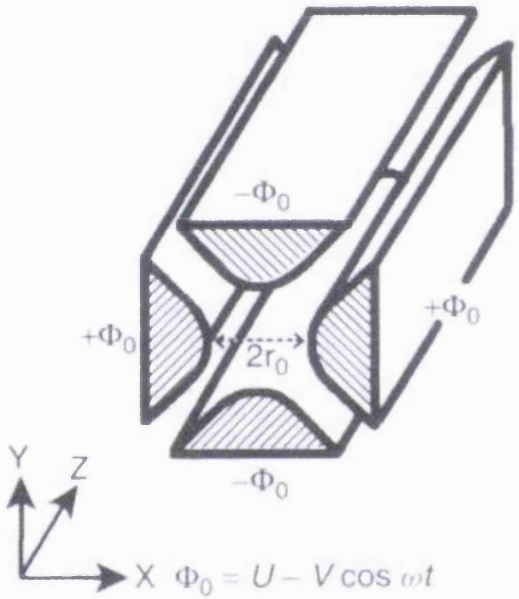


Figure 1.3: Quadrupole with hyperbolic rods at a distance of $2r_0$; the same potential (ϕ_0) is applied to opposite rods. ϕ_0 is composed of a direct potential, U, and a radio frequency potential, $V\cos\omega t$ (ω = angular frequency, V=zero to peak amplitude of r.f. field) (from de Hoffmann, p 66 [38]).

In practice, the hyperbolic rods are commonly replaced by cylinders, which are placed precisely at the corners of a square with a distance between the surfaces of twice the inside radius of the array ($2r_0$) (Figure 1.3). Opposite rods are connected with each other. For a given quadrupole instrument the inside radius r_0, is constant as well as the radio frequency (by maintaining ω); only U (direct voltage potential) and V (radio frequency component-potential) are variable. Depending on the radio frequency applied, only ions with a distinct *m/z* range can pass through the quadrupole, the others

oscillate increasingly and eventually hit the rods. That makes the quadrupole a scanning or filtering mass analyzer. The separation does not depend on the kinetic energy of the ions, but only the mass to charge ratio. In the scanning mode U is changed linearly, so it is possible to observe ions of different masses successively. It is also possible to set U=0 and apply only a radio frequency, so that the quadrupole is not able to filter ions anymore. This kind of setting is applied when several quadrupoles are arranged in tandem to just transmit the ions and focus them before entering the next mass-analysing quadrupole.

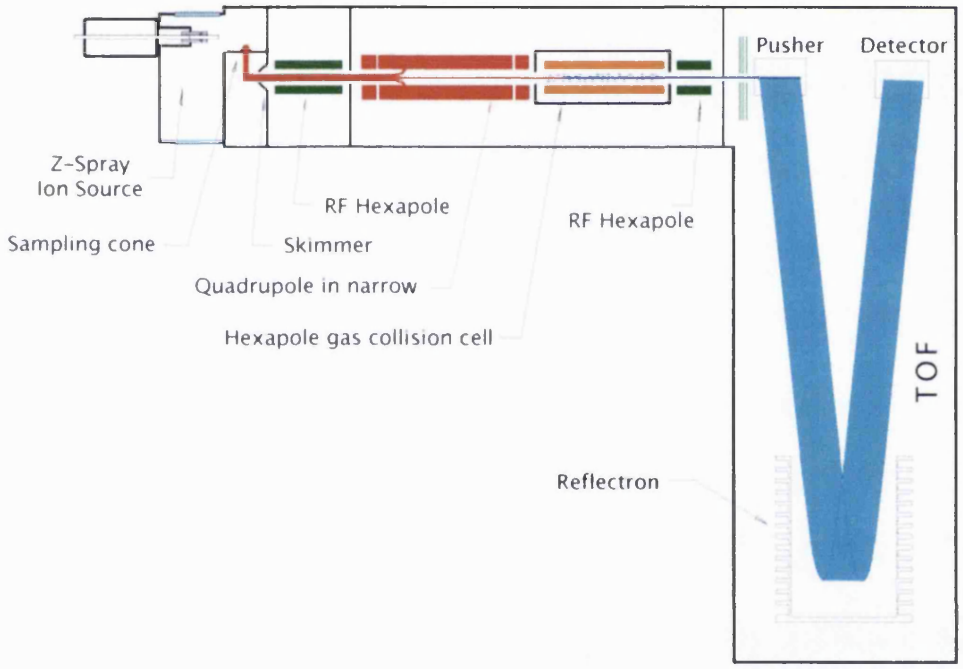


Figure 1.4: Schematic drawing of the Q-TOF instrument used in this work. Ions are introduced via the Z-spray ion source and transferred to the collision cell where ions are fragmented. The fragments are then transferred into the high vacuum time-of-flight chamber where they are separated according to their flight time (figure from waters.com).

The QTOF is a hybrid instrument combining quadrupole and time-of-flight mass analysers (see below for time-of-flight analysers). In the QTOF used in this work (Figure 1.4), two multipoles are set in series, where the second one is actually a hexapole functioning as a collision cell (for MS/MS). Thereafter is another multipole (hexapole) to focus the ions prior to entering the adjacent TOF. In contrast to ion filter mass

analysers, time-of-flight mass analysers cannot work with a continuous ion beam. Therefore, the beam has to be separated into distinct ion packages, which are then accelerated and analysed according to their flight time. The pusher accumulates ions from the ion beam and accelerates the packets into the flight tube.

A further often used combination is a triple quadrupole or 'triple quad', an alignment of three quadrupoles in a row, where the first and last quadrupole are functioning as scanning devices and the middle one is used as a collision cell.

Quadrupole Ion trap

Wolfgang Paul and Helmut Steinwedel introduced the quadrupole ion trap, often referred to as the ion trap, shortly after the establishment of the quadrupole in 1960 [52]. In the 1980s Stafford *et al.* [53] from the Finnigan company developed a commercialized ion trap detector. Hans Dehmelt and Wolfgang Paul received the Nobel Prize in Physics in 1989 for their contributions to the development of the ion trap technique. As the name indicates, ions are trapped in this mass analyzer and detected according to their *m/z* ratio. The ion trap can be seen as a derivative of the quadrupole bending around itself thereby forming a closed loop. The ion trap consists of three electrodes, a circular doughnut-shaped ring electrode and two outer end cap electrodes. Ions of all *m/z* ratios enter the trap together and are confined by the rf field, resulting from the application of an rf voltage to the ring electrode only. With oscillating fields ions are kept in the centre of the ion trap i.e. the centre of the ring electrode. As the ions in the trap collide with each other, and gain more kinetic energy their trajectories can expand ultimately becoming unstable. To avoid this, Stafford [53] introduced helium as a cooling gas to remove this excess energy, which was a major improvement of the technique. As the oscillating frequency of an ion is dependent on its *m/z* ratio, this can be used to separate and detect ions out in the trap. The rf voltage on the ring electrodes is ramped and so the frequencies of the ion oscillations also increase and end in unstable trajectories and their expulsion out of the trap. The expelled ions can then be detected and are shown in the mass spectrum. Different *m/z* values of the ions lead to different exit-times. The ion trap is often referred to as a three-dimensional quadrupole in contrast to the two-dimensional quadrupole.

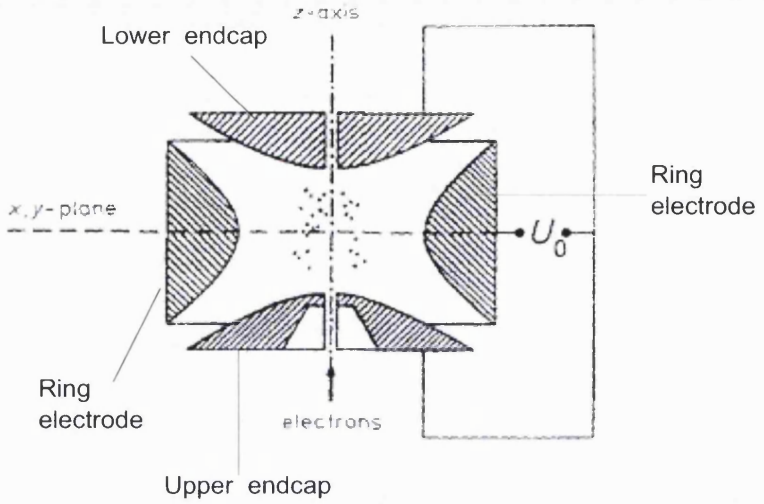


Figure 1.5: Schematic view of an ion trap, adapted from [52]. The graphic shows the two endcap electrodes and the doughnut shaped central ring electrode. The ions are shown on trajectories in the centre of the trap.

By applying a specific rf voltage to the end caps it is possible to control the ejection of all ions out of the trap except those with a selected *m/z*, which can then be submitted to further MS/MS analysis. For this purpose, a resonating frequency is applied to the end caps, which corresponds to the resonating frequency of the ion to be analysed and is also slightly lower than that needed to eject the ion. Thereby, the ion oscillates even more; it collides with the helium gas and fragments into ions with lower *m/z*, which are kept in the centre of the trap by the helium buffer gas. Then the rf voltage on the ring electrode is increased, so the fragment ions can be ejected from the trap. If this can be repeated and further product ions selected for fragmentation, MS^n analysis is performed; the application of MS^n cycles up to MS^{12} was reported in 1990.

Ion traps are used for many proteomics experiments, as they potentially offer a high mass resolution, mass range and sensitivity [54]. Due to the restricted number of ions which can be stored in the ion trap they may show a poor trapping efficiency. A newer development, the linear ion trap, circumvents this issue, by providing a bigger volume for ion storage and thus offers better mass accuracy as well as increased sensitivity and resolution [55].

Time of Flight

Whereas ions are separated in quadrupole mass analysers and ion traps according to the stability of their trajectory in a changing electric field, time-of-flight (TOF) mass analysers employ a different approach. Based on the concept of Stephens from 1946, ions are accelerated in the TOF by a set potential before they fly across a defined field free distance (Figure 1.6). The time an ion needs to reach the detector is measured and by this the mass to charge ratio is determined. Since the ion has a kinetic energy when it leaves the acceleration cell there is a correlation between its mass and the time of flight [56]. The kinetic energy can be described as $E_k = 1/2\ mv^2$ (m: ion mass; v: velocity) which is equivalent to the electric energy applied to accelerate the ions, $E_k = zeU_{acc}$, (U_{acc}: accelerating voltage, e: electron charge (e= 1.6×10^{-19} Coulomb; z: number of charges). The time an ion needs to fly over the distance d is determined as $t = d/v$ (d: distance travelled). Combining both equations allows the determination of the flight time as $t = \frac{d}{\sqrt{2eU}} \sqrt{m/z}$. Therefore, with constant distance and acceleration voltage, the time an ion needs to pass the analyser is only dependent on its mass to charge ratio (*m/z*). So the lower the *mass to charge* ratio is, the faster the ion reaches the detector [57]. This analyser cannot be utilised like the quadrupole as a scanning device but has the advantage of being capable of parallel analysis of the ions. Furthermore, it has no theoretical mass limit and can provide a higher resolution as the resolution only depends on the time-resolution of the detector.

The TOF in the QTOF used is equipped with a reflectron (Figure 1.4), which acts like an ion mirror and reflects the ions back through the flight tube to the detector [58]. The more energetic an ion, the deeper it penetrates the retarding field of the reflectron before being reflected. The advantage of the reflectron can be seen in the prolongation of the flight distance and also in correcting for small differences in kinetic energy during the ionisation event. In this way ions of the same *m/z* but slightly different initial velocities can reach the detector together. This is because ions with a slightly high kinetic energy penetrate the reflectron more deeply and exit at the same time as an ion of slightly lower kinetic energy but same *m/z*.

A further advantage of combining the TOF with the quadrupole is that the TOF has no theoretical upper mass limit, which makes it a preferred analyser coupled with soft ionisation methods. ESI enables a continuous flow of ions into the quadrupole mass analyser. After the ions have passed the quadrupole, they are focused in the hexapole and then accelerated orthogonally into the time of flight analyser. Thus, this combination of analysers makes it possible to select masses (quadrupole), perform collision experiments with selected ions and analyse the product ions in one experiment. This approach to tandem mass spectrometry shows a profit in high sensitivity, mass resolution and mass accuracy [57].

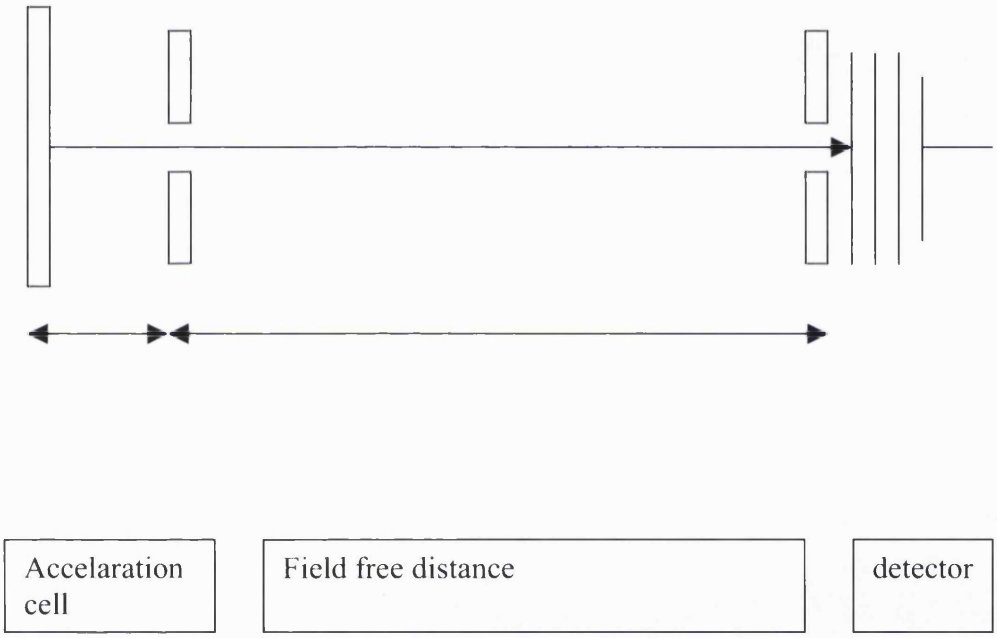


Figure 1.6: Schematic of a time-of-flight mass analyser. Ions are introduced into an acceleration cell, where they are accelerated by the application of a electric field E_{acc}. During acceleration, potential energy becomes kinetic energy. They then travel through a field free distance and the time needed to reach the detector is dependent on the m/z of the ions.

1.2.3 Peptide sequencing in tandem mass spectrometry

As already mentioned, it is possible to perform MS/MS analysis on peptides using tandem mass spectrometers, i.e. the combination of two analysers, to combine selection of a defined parent ion and the further collision in the gas phase and fragmentation into

its product ions, also called daughter ions. This is the common procedure for peptide sequencing by mass spectrometry. The fragmentation of peptides by mass spectrometry is well documented [59]. Usually, low energy collisional activation instruments as used here i. e. QTOF or the iontrap, the (multi)protonated peptides fragment mainly along their backbone. The fragmentation of most (multi)protonated peptides is dependent on their size as well as the energy introduced and number of charges, and fragmentation often requires a proton rearranged to the cleavage site [60]. This proton can be mobilised (e. g. in collisionally activated dissociation or CAD) and relocated from e. g. a basic side chain to the backbone. For peptides containing a very basic residue one can therefore expect that more energy is required to relocate a proton to a position in the backbone and induce fragmentation there. For arginine-containing peptides energy to relocate to the backbone is considered to be the highest, followed by lysine-containing peptides [61]. There are different ways in which the bonds on the backbone can be broken. Important for all observations is, that only charged fragments will be detected in the mass spectrometer, so it is important to know where a charge remains after backbone fragmentation. According to that, the fragments are named 'a', 'b' and 'c'-ions, depending on the cleaved bond, if the charge is retained on the N-terminal fragment. If the charge is found on the C-terminus, the corresponding ions are labelled 'x', 'y' and 'z' ions (Figure 1.7). A subscript describes the number of residues, so y_2 signifies the y-ion with two amino acid residues. In low collision energy CAD experiments the mostly observed fragments are y- and b-ions, which are obtained after cleavage at the CO-NH bond. Cleavage at the CH-CO bound results in a- and x- ions, while c- and z- ions are derived from cleavage at the NH-CH bond. Due to the cleavage site of y- and b-ions, these fragments can help to identify the amino acid sequence of a peptide: the mass difference between two adjacent b-ions (or two y-ions) shows the mass of the intervening amino acid residue. The more y- or b-ions that are detected, the more secure is the identification.

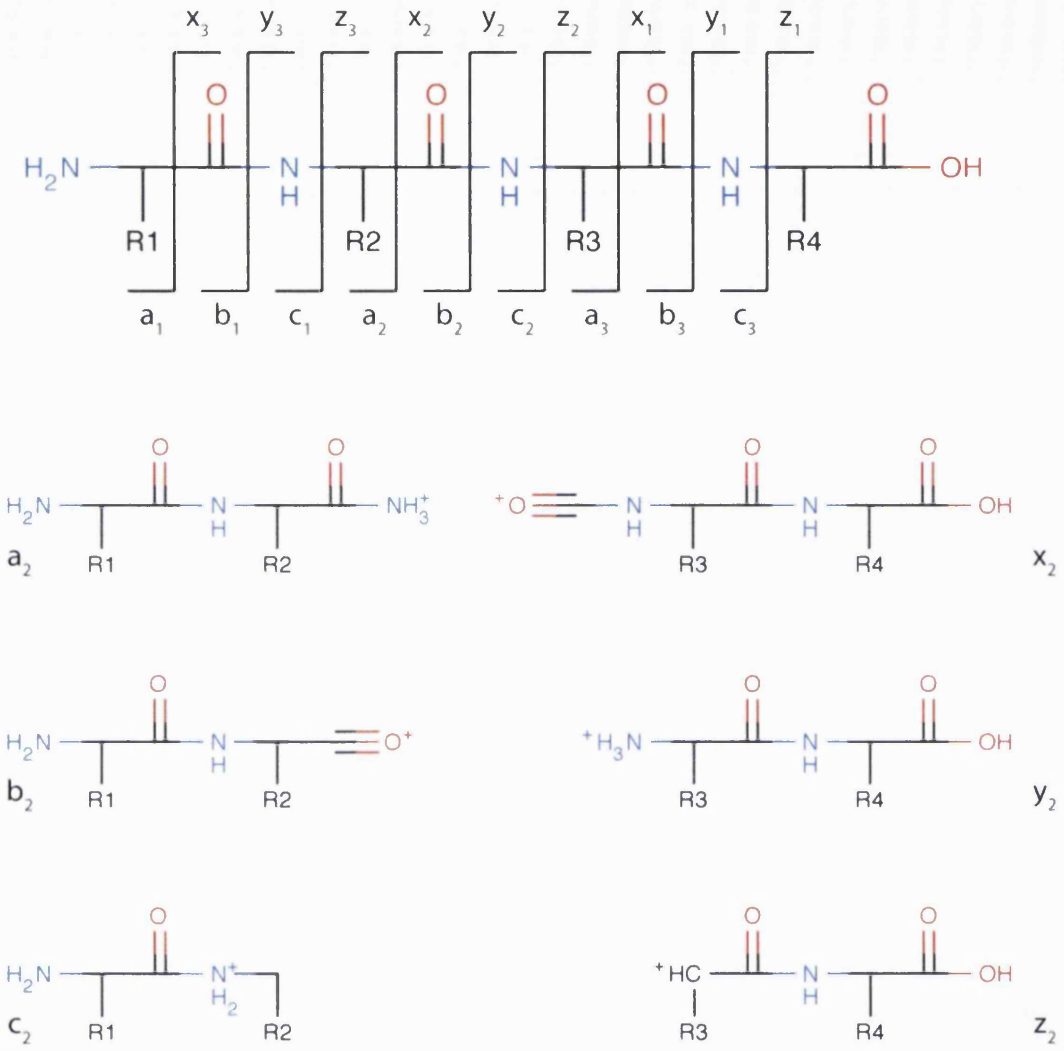


Figure 1.7: Collision induced dissociation (CID) induced fragmentation pattern of peptides into smaller fragments along the peptide backbone. Ions where the charge remains on the N-terminus are called a-, b- and c-ions whereas ions with the charge on the C-terminus are called x-, y- and z-ions.

1.2.4 LC-MS

A convenient method to separate samples prior to mass spectrometry is found in liquid chromatography (LC). Coupled with mass spectrometry, mass spectra can be acquired while the samples are eluted. This works very well in connection with electrospray ionisation, where the sample is required to be in a solution.

In this work reversed phase high performance liquid chromatography (RP-HPLC) is used. In this technique substances are separated according to their polarity

and their distribution between a mobile and a stationary phase. A non-polar column ("reversed" hydrophobic stationary phase) is loaded with sample, and then mobile phases of decreasing polarity are used to elute the sample from the column. The components of the sample of low polarity will be retained on the column. The addition of mobile phase elutes first the most polar substances, which show little interaction with the column. The higher the proportion of non-polar solvent in the mobile phase, the more of non-polar substances of the sample mixture will elute. Usually the columns in RP-HPLC consist of a chemically modified silica stationary phase. Widely used are columns with bonded C_{18} alkyl groups creating a non-polar phase (liquid chromatography-mass spectrometry) (Figure 1.8). The analytical columns in the experiments presented in this work had an inner diameter (ID) of 75 µm and were part of a nano-LC system. Capillary LC refers to columns with IDs between 100 and 300 µm, inner diameters between 300 and 1000 µm are referred to as micro LC. Using a miniaturised method such as nano LC also requires appropriate patience and care; over-loading of the column, high back-pressure, blocking through micro particles or leakages are common incidents [62] resulting in delays and potentially loss of sample. Due to the fact that available sample of biological material or their preparations are rather small, minaturised LC-methods are a preferred application in proteomics.

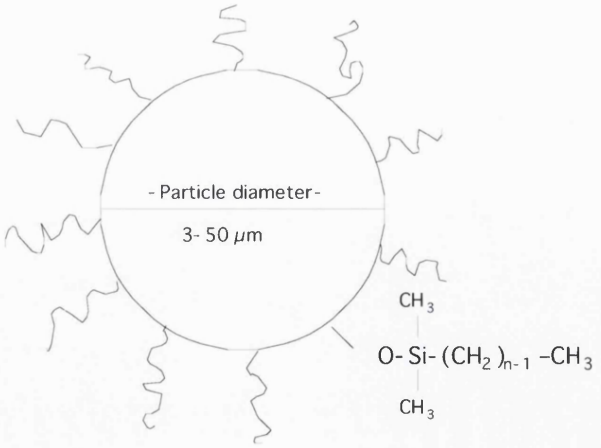


Figure 1.8: Schematic view of a Silica based packing particle with n-alkyl chains covalently bound. Interactions of the alkyl chains with the peptides in solution and polarity of the mobile phase determine the retention of the analyte on the column. Commonly C_4, C_8 and C_{18} chains are used.

1.2.5 Protein identification

As shown in Figure 1.9, the typical workflow in a proteomics experiment includes the analysis of the acquired data. In the case of the electrospray – MS/MS approach this means that tryptic peptides were fragmented and the experimentally derived fragments are compared to *in-silico* calculated fragments from protein databases. The assignment of theoretical to experimental data is supported by specially designed software providing search algorithms [63], most commonly used are Mascot [64], Sequest [65] or MS-Tag [66]. The Mascot search engine compares the theoretically calculated fragment ion masses for an *in-silico* generated proteolytic peptide with the peaks in the experimental spectrum. A probability-based score will be given for each assignment, indicating the significance of the hit. The experimental/instrumental information has to be considered as well, in particular mass tolerances, which have to be supplied as search parameters for parent-peptide ions and product ions in MS/MS fragmentation. Mascot scoring is based on the Mowse (*mo*lecular *w*eight *se*arch) scoring algorithm [67], and calculations are based on the assumption that a hit is a random event. The less random a peptide hit, the better the assignment between theoretical and experimental data was. Therefore scores are presented as $-10*\log_{10}(P)$, where P is the probability that a hit was a random event. So if the probability for a random event is low, e.g. 10^{-30}, the score would be 300. High scores are desirable for a good peptide hit. With growing databases and bigger number of proteins identified in one experiment, there is a need to validate the assigned peptide hits regarding the possibility of retrieving false positive hits. For this reason so-called decoy searches are performed, which use an additional search of the same experimental data using the same search parameters against a randomised or reversed database [68]. Ideally there should be no assignments using a decoy database. However assigned matches using the decoy database can be an estimate for false positive hits. This was part of the guidelines proposed in a workshop in 2005 by the Editors of Molecular and Cellular Proteomics (http://www.mcponline.org/misc/ParisReport_Final.shtml) ([69]). As more and more proteomics data is published, it becomes difficult to compare various studies, so by defining these guidelines, a basis for the presentation of proteomics data is created. Issues such as information about experimental conditions, number of peptides, scores, databases used, handling of modifications or quantitative data are considered. There also have been attempts to simplify mass spectrometry raw data, so redundant or

low quality spectra are removed before data is submitted to a database search and thus simplifying the search [70, 71]. Also the results from a database search can be very complex and need sorting and validation of the assigned peptide hits, so new software-algorithms are emerging [72-74].

Nevertheless it has to be considered that the outcome of a database search is only as good as the experimental data, in fact algorithms are computationally intensive and require high-quality data [63]. As long as a protein is in a database, it can be identified. If this protein has not been identified before, then it might get overseen in the database search, as well as unidentified modifications.

1.2.6 Outlook

Ongoing research in proteomics and its instrumentations is always looking for improvements. In the area of mass spectrometry new instruments such as the linear ion trap [75, 76] and the orbitrap [76] have improved resolution and sensitivity of detection, the Orbitrap coming close to the capabilities of Fourier transform ion cyclotron resonance (FTICR) instruments [76, 77]. Also further effort is put into fragmentation techniques, such as ETD [78], a recently established alternative to CAD and CID. As the problems change, the methods have to adapt. Proteomics is aiming beyond the simple identification of proteins: quantification has to be achieved, posttranslational modifications detected, distinguished and sensitivity increased to overcome the large dynamic range of protein abundance in biological samples [79, 80], an issue evident in this work. It is well known that the dynamic range of biological samples can vary by more than ten orders of magnitude, so this is definitely a challenge for proteomics techniques [81]. Not to be underestimated however, is the challenge presented in handling all the acquired data. As the results achieved by different groups need to be compared and combined, identification, quantification and comparison of several thousand proteins in one experiment require thorough design of data- analysis and storage, so bioinformatics and proteomics will be inseparable.

1.3 Mass spectrometry based proteomics

Developments in mass spectrometry have preceded developments of proteomics. In fact, the introduction of the so-called 'soft' ionisation techniques electrospray (ESI) [37, 40, 82, 83]

and matrix assisted laser desorption ionisation (MALDI) [36, 84] enabled the application of mass spectrometry to (large) biomolecules; in 2002, Koichi Tanaka [85] and John B. Fenn were awarded with the Nobel prize in chemistry for their achievements in developing these techniques.

Mass spectrometry (MS) [18, 86] is the detection method of choice for proteomic experiments because of its sensitivity, reliability and the possibility for high throughput screening. A common procedure in a proteomic experiment is shown in Figure 1.9 and includes the separation of a protein mixture by gel electrophoresis (either one dimensional or two dimensional), proteolytic digestion and then fragmentation and detection of the peptides via mass spectrometry. Using gel electrophoresis and MS, it is possible to detect a large number of proteins in a complex mixture. The digestion step is introduced to reduce the chain length of the protein and facilitate identification. As some proteolytic enzymes cleave only at specific amino acids of the polypeptide chain, for example trypsin acts at the C-terminus side of arginine and lysine residues only (except when the next amino acid is proline in which case cleavage is inhibited) [87], it is possible to digest a protein theoretically *in silico* and compare this with the peptides generated by proteolysis and analysed by tandem MS (see 1.2.2). Fragmentation of peptides provides information about their amino acid sequence and databases, containing amino acid sequence information, can be searched with this information to identify precursor proteins. While performing a database search, parameters have to be chosen taking into account the experimental conditions. Thus the results obtained by a database search are probability based and represent the conformity of experimental and theoretical data.

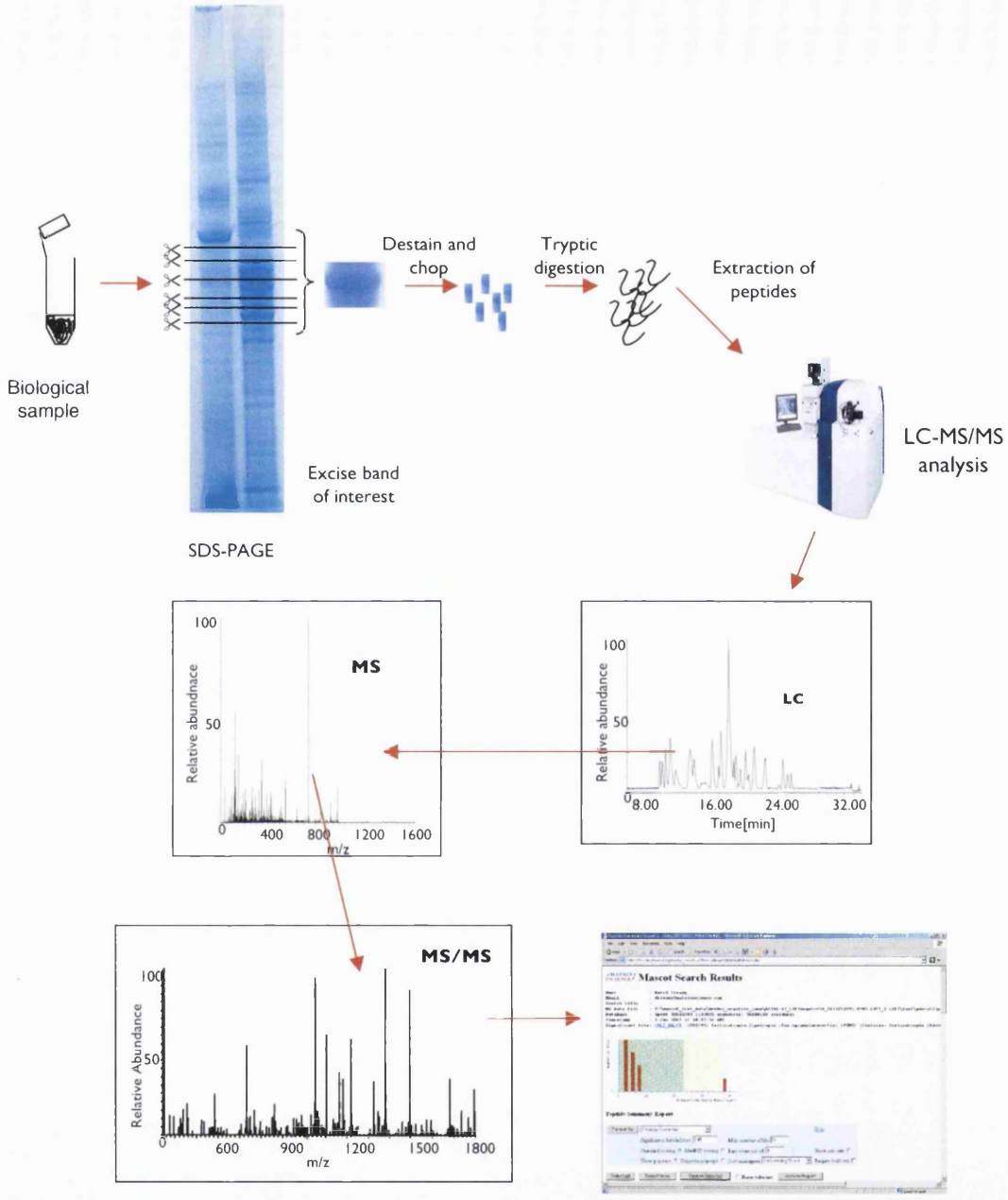


Figure 1.9: Workflow of a mass spectrometry based proteomics experiment: the sample is separated, often on SDS-PAGE; areas of interest are cut out and chopped into smaller pieces for digestion. Prior to mass spectrometry analysis, the extracted peptides are run on a LC-column for further separation. Full-scan MS is performed on the eluted peptides, followed by MS/MS analysis. The resulting raw data is then submitted to database search and experimental data is assigned to fragments in the database.

The analysis of peptides derived through protein-digests is often referred to as bottom-up sequencing [88] and proteolysis is a commonly used way to create smaller molecules for mass spectrometry analysis. The sequencing of "intact-protein" ions in the gas phase is in contrast called top-down sequencing [88], but we will concentrate here on the first approach. A variety of studies have been undertaken so far performing tandem mass spectrometry of peptides, not only for identification but also quantification, structural and protein-protein interaction analysis and post translational modification studies [2] using both ESI and MALDI [89].

One interesting application is the use of proteomic tools in the discovery of biomarkers [90-92]. As the early detection of any disease is desirable for early treatment, there is a need to identify unique markers that can be used clinically. Biological fluids such as serum, urine, saliva as well as tissue from biopsies are the main sources of possible biomarker-proteins. Diseases like Alzheimer's disease [93] and asthma [94] are an example for areas of focus, as are other diseases like cancer, e.g. ovarian cancer, breast cancer, prostate cancer etc. [95, 96]. An attempt to find biomarkers coincides with the fundamental challenge to understand the origin of a disease. To support this attempt comparative studies of diseased and normal tissue are carried out [7, 97]. As any protein, whose expression is changed in a particular physiological state, could function as a biomarker in theory, a vast number of protein biomarkers are conceivable. Protein markers are not only useful to investigate differences in healthy and diseased tissues or diagnose an illness, they can also be employed to monitor a drug therapy and report the effects of the drug on the levels of proteins as well as the efficacy of a treatment [98].

For most of the samples chosen for analysis, several separation steps are required before the sample can be submitted to MS-analysis [91]. A frequently applied technique for the analysis of tissue-based samples is subcellular fractionation into subgroups of the cell [99-101]. Of special interest is the subcellular microsomal fraction, which consists of membranes of the endoplasmic reticulum. The subject of this work, namely metabolising enzymes are known to be present in this fraction, particularly cytochrome P450s [102]. Briefly, homogenised tissues or other biological samples are fractionated according to their density using e.g. ultra centrifugation [103] to decrease the complexity of biological samples [80], as the individual fractions are less complex and analysis can be performed on subgroups of proteins. However, it has to be taken into account that

events in the cell are dynamic and cannot be categorized into static compartments. Proteins are transported around in the cell, from one compartment to another, so in order to achieve a better overview of the actual physiological proceedings one has to look at different states of the cell or tissue type [104, 105]. It is often pointed out that sub-celluar fractionation is an appropriate method to enrich membrane proteins, which about one third of all the genes encode for [106, 107], in particular since these proteins are often difficult to isolate. There have been several approaches using mass spectrometry based proteomics to address the issue of better identification of membrane proteins as well as increase the dynamic range of protein detection. One approach is to use a gel-free method where 2D-LC, (multidimensional separation), is used prior to MS/MS and is suggested to be able to identify more low abundant as well as membrane proteins [108]. In this method, a strong cation exchange column is coupled to a reversed phase column and tandem MS analysis is followed by a database search [109]. It has been recently effectively applied for the proteomic studies of human cortex [110], mouse heart [111] and human plasma [112]. However, no data is available on the application of this method specifically to CYP proteins.

A further application of proteomics can be found in drug discovery, where proteins play an important role as drug targets [113]. Also a lot of therapeutically used compounds are proteins, such as insulin [114], erythropoetin [115], and various antibodies used for anti-cancer therapy such as rituximab against Non-Hodgkin lymphoma or bevacizumab (Avastin®) for the treatment of metastatic colorectal cancer [116].

In this work we have chosen a one dimensional gel electrophoresis separation coupled to a LC-MS/MS approach to investigate microsomal proteins including CYPs in lung tissue.

1.4 Quantitative proteomics

Quantitative and qualitative proteomics are not separate fields, rather they complement each other, especially as quantitative methods are used more and more to explain or support results from identification studies [117] (Figure 1.10). The presence or absence of a protein in a biological sample needs to be further interpreted by annotating this data with quantitative evidence as well. Anderson [118] pointed this out using the example of cardiovascular disease biomarkers, where 177 potential biomarkers were found in a

search of over 2000 published papers. Due to the lack of quantitative information, he suggests a more targeted approach, especially paying more attention to factors such as specificity and selectivity of a defined biomarker. Here an overlap with quantification methods is desirable as results from such studies could allow conclusions about the relevance of previously suggested biomarkers. Prior to the general application of mass spectrometry technology, proteins were predominantly quantified using immuno-chemical based techniques, such as Western blot [119] or enzyme-linked immunosorbent assay (ELISA) [120]. In addition to the requirement of a specific antibody for each protein or protein fragment, the success of these methods is dependent on the efficiency and selectivity of the antibody and whether there are any cross-reactions with other proteins e. g. precursors or degradation products of the protein under investigation. [121]. This can be avoided by quantifying proteins via mass spectrometry methods [122] which does not require antibodies and can identify proteins by their amino acid sequence.

The development of quantitative mass spectrometry based proteomics initially went along with the use of two dimensional gel electrophoresis analysis [26]. 2D-gels are currently used to separate complex samples. Protein spots are stained with Coommassie blue, silver, or fluorescent dyes, and gel spots are cut out, digested and analysed using mass spectrometry techniques [123]. In this way a quantitative comparison of two samples can be achieved by comparing the intensities of the stains. However, the approach here is rather a global overview of high and lower abundance proteins comparing two gels. A further development to use 2D gels in a more quantitative mode and to circumvent inter-gel variations, is the two dimensional difference gel electrophoresis (2D DIGE) [124, 125]. Prior to gel electrophoresis, samples are labelled with different fluorescently resolvable dyes (CyDyes) [126]. Thus, samples run on the same gel can be compared visually. It is also now possible to use a standard, which consists of a mixture of all samples, labelled with one dye and run on the same gel enabling normalisation of each sample against the internal standard and thus improving accuracy of quantifications made [127]. However, problems in two-dimensional gel electrophoresis concerning the difficulties to detect hydrophobic proteins and detect very low abundant proteins have not been entirely resolved [128].

Although mass spectrometry is commonly used for the quantification of 'small molecules' such as drugs or metabolites, its performance is considered poor for protein

quantification, as the relationship between sample concentration and its signal intensity has not been entirely explained [18]. Therefore, auxiliary methods have to be used to circumvent these uncertainties in mass spectrometry. For quantitative studies stable isotopes have become a very suitable way of introducing a measurable change of molecular weight (greatly) without changing the physicochemical properties of the analytes [129]. As two compounds only differ in their molecular weight due to the mass difference of the isotopes, they show similar physicochemical properties, such as retention time in liquid chromatography, and the difference can only be seen in the mass shift in the mass spectrum acquired. Eventually in the mass spectrometric analysis, the abundance of the signal derived from the detection of stable isotope and native structures are compared for quantification [130]. There are various ways to introduce a stable isotope into the sample: metabolically, chemically, enzyme catalysed or via direct addition of stable isotope labelled internal standard to the sample [131]. The metabolic introduction is performed by using stable isotope enriched media to grow cell cultures. Best known is the SILAC technique, the 'stable isotope labelling with amino acids in cell culture', by Ong *et al.* [132]. Cells are grown on a medium (Figure 1.10), which contains one amino acid exclusively in a stable isotope form (e.g. deuterated leucine) and is thus incorporated into all proteins produced by the cell. Two pools of cell cultures can be labelled differently, combined before lysis, processed together and via mass spectrometry the ratios of labelled and unlabelled peptides can be compared for relative quantification [132]. The advantage of this method is that the cells are processed together as they are mixed before lysis and all following steps of sample preparation such as homogenisation, electrophoresis, or chromatography are done on one combined sample, sample losses will occur for both analytes to the same extent. The method has even been extended to growing of *Escherichia coli* and yeast on ^{15}N-enriched medium and feeding this to small organisms such as the fruit fly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans* [133]. Ishihma *et al.* [134] used the SILAC technique to express CDITs (culture derived isotope tags); cells, cultured in a stable isotope enriched medium are mixed with tissue and processed together, thus serving as an internal standard. Even though this seems a neat approach to introduce stable isotopes, it is restricted to small organisms, and for quantitative analysis of the human organism different techniques are

preferred. It has also been observed that some cells tend not grow on SILAC medium [129].

Introducing stable isotopes can also be achieved through chemical modification of proteins. This way, tagged peptides can be selected and isolated for quantitative analysis. There are diverse groups in proteins, which can be tagged, such as amino groups [135], sulfhydryl groups [136-138], phosphate ester groups [139, 140], the active site of serine hydrolase [141] or N-linked carbohydrates [142]. Currently the most often used techniques are ICAT (isotope coded affinity tags) and ITRAQ (isobaric tag for relative and absolute quantification) (Figure 1.10). ICAT was introduced by Gygi and Aebersold in 1999 [136] and is based on the labelling of proteins at their cysteine residues with a cysteine reactive part of the ICAT reagent. To this reactive group a polyether linker region is connected, carrying initially eight deuteriums, now with nine ^{13}C, for the mass difference. A biotin group is attached as well, which is used to separate the labelled peptides via an avidin column. Protein samples are usually denatured, reduced and the reactive group of the ICAT reagent binds to the cysteine in the proteins. After digestion with trypsin, the labelled peptides are separated by affinity chromatography on an avidin column. In the LC-MS/MS analysis of the peptides, the MS spectra show the abundance of the heavy and light peptides and ratios can be calculated. MS/MS is then used to identify proteins with the information obtained from their fragmentation. A further improvement to the initial reagent, an acid cleavable ICAT reagent, has been intoduced, which enables the removal of the biotin-tag prior to MS- analysis [137]. Using this method it is possible to obtain relative quantification data by comparing a sample treated with heavy isotopes and one with light isotope incorporation. Using affinity chromatography to separate the tagged peptides also results in an enrichment of labelled samples. However, this method is only applicable with cysteine-containing proteins and if denaturation and reduction are incomplete, some cysteine-residues may remain unlabelled. Using cysteine to tag proteins also results in a bias against proteins with fewer cysteine residues and some proteins might not be detected [122]. The ICAT method, as well as ITRAQ, are patented and kits are available for purchase. Ross *et al.* [143] developed the ITRAQ technique, which makes it possible to compare more than two samples. The N-termini and lysine residues of peptides derived from tryptic digestions are labelled with one of four different isobaric tags. Once subjected to MS/MS, the tags

are fragmented and each relieves a reporter ion of different molecular weight (114–117 Da), and the intensities of these reporter ions, each corresponding to the individual sample they have been added to, are used to perform relative quantification studies. Each tag consists of the reporter ion, a balance group, which have a combined molecular weight of 145 Da (corresponding to 114–117 Da for each reporter group and 28–31 Da for each of the balance group) and a peptide reactive group for binding to the amine groups in peptides [144]. Thus, the MS/MS analysis offers identification of peptides and simultaneously the comparison of the reporter ions for quantification. The low mass of the reporter ions is chosen not to interfere with higher mass regions in the MS/MS spectra, where peptide fragmentation peaks are to be expected [144]. Recently, four additional reporter ions were introduced, so it is possible to compare eight samples in one analysis [145].

A further possibility to introduce stable isotopes is to incorporate ^{18}O during an enzymatic digest with trypsin, labelling the carboxyl-terminus of the peptide [146, 147]. Using proteolytic labelling, trypsin is used to exchange the two carboxyl-oxygens with ^{18}O [148]. The sample is initially digested into smaller peptides using proteolytic enzymes. In a second step, the digest is treated with trypsin to incorporate the isotope label: one sample is digested in $H_2{}^{16}O$-solvent and the other sample in $H_2{}^{18}O$-solvent. The intensities of $^{16}O/^{18}O$ pairs are compared in MS/MS and relative quantification performed. The advantage compared with chemical labelling techniques is that each peptide can be labelled and no functional groups are preferred. However the mass difference is only 4 Da which requires good resolution in the mass detector [149]. So far, there are still cases of incomplete incorporation of ^{18}O reported, especially when complex samples containing many peptides are quantified [150].

For all the labelling methods to be used successfully, it is necessary that complete labelling or chemical exchange reactions have taken place and minimal incomplete labelling or contamination has occurred [151]. Furthermore, the labelling methods lack the ability to quantify one sample on its own, independent of the comparison with another sample.

In order to pursue an absolute quantification, internal standards need to be applied. In case of quantification of a protein, either a stable isotope labelled protein or a peptide are suitable. As the synthesis of a whole protein is difficult and expensive, the

majority of internal standards used are peptides, which are unique in their sequence for the protein of interest [152]. The stable isotope dilution technique has been well established for small molecule quantification over at least 4 decades, and over 20 years ago Desiderio started to apply this principle to the analysis of peptides and polypeptides and used stable isotope labelled peptide standards for the quantifcation of endogenous leucine enkephaline and methionine enkephaline [153]. In this technique, the area or intensity ratio between the analyte and a stable-isotope labelled internal standard is used for quantification. Insulin and apolipoprotein A-I, both physiological important proteins, have been quantified using stable isotope dilution over ten years ago [121, 154]. Further in-solution quantification of an integral membrane protein was performed by Barnidge *et al.* [155] in 2003 and Gerber *et al.* [156] extended this technique to the quantification of in-gel digested proteins and thereby also introduced the term AQUA, **a**bsolute **qua**ntification. The selected heavy-isotope labelled peptide represents a part of the protein sequence derived by tryptic digest, so the peptide is introduced into the sample during tryptic digestion and the corresponding non-labelled (light) form is produced through enzymatic hydrolysis of the protein. Due to the fact that both peptides only differ in their mass, because of the heavy-stable isotope incorporation, they show the same physicochemical properties and thus can be separated on a LC-column from other compounds, as they will co-elute. The two peptides are then detectable using MS and to ensure the correct identification, a MS/MS fragmentation can be performed. For quantification, multiple reaction monitoring (MRM) is used in which the fragmentation of one ion into a specific fragment is monitored [157], allowing the reliable quantification of the respective compound. In general, the area under the curve in the selected chromatograms of internal standard and native peptide are determined and used for calculation of the amount of protein present. Usually stable isotopes such as ^{2}H, ^{15}N, ^{13}C and ^{18}O are employed and usually one amino acid is labelled, e.g. in the case of leucine, six ^{13}C and one ^{15}N are incorporated, resulting in a mass shift of 7 Da. In electrospray, mainly multiple charged ions are formed, so the mass shift should not be too small in order to avoid overlapping isotope peaks in the mass spectrum. Deuterium is in general not used as the stable isotope as retention time shifts in chromatography have been experienced [158]. The incorporation of stable isotope labelled internal standards is a promising attempt towards an absolute quantification of

proteins [152] and as the internal standard is added during digestion it can compensate for some losses made during and after digestion. The principle, that peptides unique to only one protein are used, increases the selectiveness of this method. Also the application of mass spectrometric MRM analysis can ensure a sensitive detection. If the position of labelling is selected to be in the fragment which is used for MRM, it increases selectivity [152]. The so called 'AQUA' method has also been used for the quantification of post translational modified proteins e.g. by using a stable isotope labelled form of phophorylated and nonphosphorylated human separase to measure the extent of phosphorylation [156, 159].

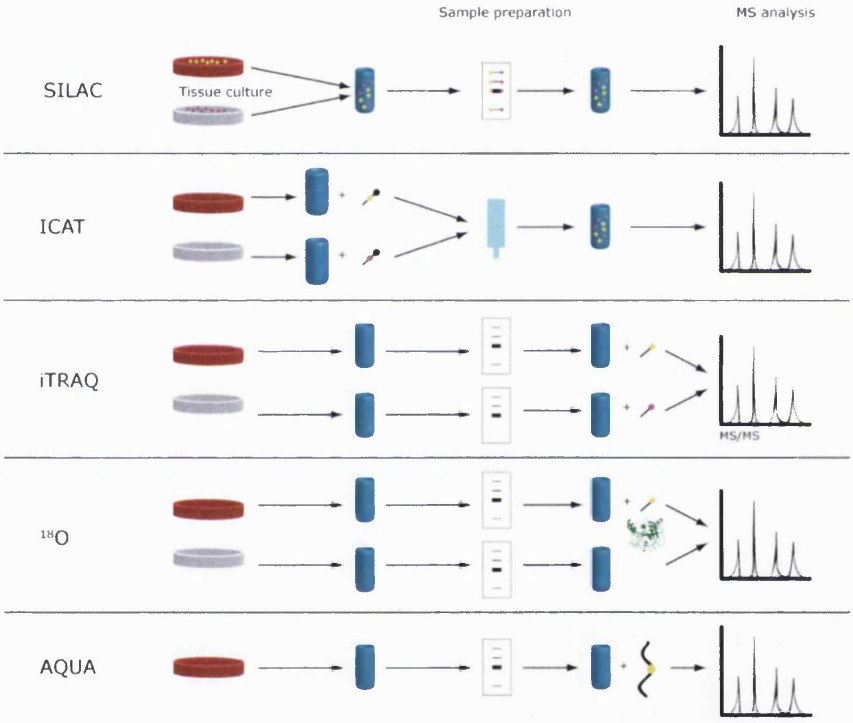


Figure 1.10: Different approaches to quantify proteins using mass spectrometry; SILAC, ICAT, ITRAQ and ^{18}O only allow relative quantification by comparing a sample (red) with a control (grey); AQUA allows absolute quantification by introducing an internal standard. SILAC (stable isotope labelling with amino acids in cell culture) introduces isotope labels by growing cells on labelled medium; the cells are harvested and processed together, the different isotope labels allow a separation by mass spectrometry and a comparison of intensities. ICAT (isotope coded affinity tags) labels are added to cysteine residues and contain isotope labels and a biotin tag for purification. In ITRAQ (isotope tags for relative and absolute quantification) and ^{18}O labelling, the label is introduced during or post proteolytic digestion. In AQUA (absolute quantification), an isotope labelled internal standard peptide (isp) is added to the digestion mixture.

However, there is yet no method that allows the quantification of proteins in one simple and straightforward step. All principles of separation, enrichment and sensitivity issues known from identification of proteins apply also for their quantification.

1.5 Xenobiotic Metabolism and Cytochrome P450s

1.5.1 Xenobiotic Metabolism

Xenobiotic metabolism takes place in all organisms and is mainly responsible for the excretion of unwanted substances from the body. Usually it results in a change in structure that leads to better water solubility and thereby enhanced excretion. There are generally two stages of xenobiotic metabolism: Phase I and II. In phase I molecules are oxidised, reduced or hydrolysed to alter the functional groups. These functional reactions are carried out by enzymes such as Cytochrome P450s (CYPs), alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH). In the second phase of xenobiotic metabolism, the functionalised molecules are conjugated to other endogenous substrates to increase their polarity. UDP-glucuronyltransferase (UDPGT), glutathione S-transferase (GST), sulfotransferase (SULT), *N*-acetyltransferase (NAT) or NAD(P)H-quinone oxidoreductase (NQO) are commonly involved in conjugation reactions [160, 161]. The main place of metabolism in the human body is the liver, but also in other organs metabolism takes place, such as lung [162], the kidneys [163], urinary bladder [164] or skin [165].

An alternative interest in metabolising activity of an individual and its tissues concerns the fact that through the process of metabolism a substance is not only transformed into a more extractable form, but it can also become more rather than less toxic. In addition to that, individuals can have different predispositions regarding if and to which extend they metabolise different substances due to genetic polymorphisms [166-168]. This can in fact be either responsible for different susceptibility to diseases e. g. the correlation between heterocyclic amines and cancer risk [169], or smoking induced lung cancer [170]. Not only the phenotype, but also factors like nutrition, lifestyle, gender and race can have an influence on the extent of metabolism in individuals [171-173]. Expected or unexpected metabolism can also determine the outcome of a drug therapy [174]. This fact is furthermore used in drug development, where a general approach of prodrug

therapy is the synthesis of a compound, which is activated in the body. This can be due to better solubility of the prodrug or in the case of very reactive and otherwise toxic materials. In cancer therapy for example, one goal is to design drugs, which are metabolised by enzymes only present in cancerous tissue [175] [176]. This way, healthy tissue will not be affected and toxic side effects decreased. To use the targeted metabolism in drug therapy, it is important to know about the presence of metabolising enzymes in various tissues, healthy as well as diseased.

1.5.2 Cytochrome P450 enzymes

1.5.2.1 General aspects

Cytochrome P450s (CYPs) are the most prominent group of phase I enzymes, presenting a superfamily of haemoproteins. Their carbon-monoxide-bound form, where CO builds a complex with Fe^{2+} of the haem group, shows absorption at 450 nm, hence the name ('P' indicates pigment, 'cyto' means cellular and 'chrome' stands for colour). CYPs have been found in every class of organism. According to David Nelson (last update of statistics in August 2007) (http://drnelson.utmem.edu/CytochromeP450.html), there are to date over 7700 different CYP sequences known, of which 2740 are found in animals, 2675 in plants, 1231 in fungi, 813 in bacteria, 226 in protists and18 in archae. For humans there are 57 CYP-genes quoted. CYPs are classed according to the family (more than 40 % amino acid sequence homology) and subfamily (more than 55 % sequence homology) to which they belong to. The family is named with a number, the subfamily with a letter and the individual CYPs with a number. There are 18 families and 43 subfamiles known in humans. CYPs were first discovered in 1962 by Omura and Sato as a new microsomal haemoprotein [177]. The years after followed the investigation of the enzymatic function of these enzymes, in the 1970s research was carried out on their purification and the development of antibodies started. From the 1980s work was done on the gene identification, function and expression, which continues today [178]. Their origin however is supposed to date back about 1.5 billion years ago, in an ancestral CYP gene expressed in a primeval prokaryotic organism, from which all other CYPs descended. It is suggested that, during evolution, CYPs have developed in plants and animals as a defence mechanism to metabolites or other decayed plant products in the environment [179, 180].

CYPs metabolise a large number of different substrates (Table 1.1) as well as drugs and other xenobiotics. Members of the CYP1, CYP2 and CYP3 families mostly perform these reactions. Other CYPs take also part in metabolism of endogenous substrates, such as during steroidogenesis and bile acid production from cholesterol as well as β–oxidation of fatty acids and prostaglandins [180, 181]. They are most common in liver, but also found in other tissue, e. g. kidney, lung, brain, testis, ovary, uterus, colon, small intestine, placenta, prostate, skin and heart [182, 183]. In the cell, they are mainly located in the endoplasmic reticulum with a few in the mitochondria [184]; in the liver 90% of CYPs are found in the microsomal fraction and 10 % in the mitochondrial fraction [182], of which CYP11A, CYP11B1, CYP11B2,CYP24, CYP27A1 and CYP27B1 are known as human mitochondrial CYP proteins. Microsomal CYP enzymes are membrane bound; they are anchored into the membrane with a hydrophobic core of about 20 to 25 residues at the *N*-terminus of the protein.

Table 1.1: Main CYPs involved in hepatic metabolism. Adapted from [185-187]. Proportion of whole liver CYP is given as percentage. Examples of enzyme substrates, inducers and inhibitors are stated; this is only a representative selection, many more substrates exist.

CYP	3A4/5	2C9	1A2	2E1	2D6	2 A6	2C19	2B6	2C8
Proportion in liver [%]	~30	~20	~15	~10	~5	< 5	< 5	< 5	< 5
Substrate	Nifedipin midazolam, Erythromycin, cyclosporin	Tolbutamide, Phenytoin, warfarin	Caffeine, Theophylline, Tacrine	Chlorzoxazone	Debrisoquine Sparteine	Coumarin	Phenytoin, Omeprazol	Bupropion	Verapamil
Inhibitor	Ketoconazole	Sulfaphenazole	Furafylline Fluwoxamine	Disulfiram	Quinidine	Methoxsalen	Fluconazol	Thiotepaticlopidine	Ketoconazol, Gemfibrozil
Inducer	Barbiturate, Rifampicin, Dexamethasone Carbamazepine	Barbiturate, Rifampicin	Omeprazol, tobacco smoke	Ethanol, Isoniazid	Non-inducible	Phenobarbital	Barbiturate, rifampicin	Barbiturates, Phenytoin Rifampicin	Rifampicin

1.5.2.2 CYP-structure and oxygenation mechanism

As the crystallisation of CYPs proves to be difficult due to their hydrophobic character, only few members of this enzyme family have been crystallised so far [188]. In 2007, two more human CYP structures were solved; Sansen *et al.* were successful in crystallising CYP1A2 [189] and Smith *et al.* could define the structure of human lung CYP2A13 [190]. Other human CYPs with known crystal structures are human CYP3A4 [191], 2C9 [192], 2C8[193], 2A6[194] and 2D6 [195]. Figure 1.11 shows the structure of human CYP1A2, indicating the secondary and tertiary structure and the embedded haem group. Common in all CYP structures is a single prosthetic haemgroup (iron protoporphyrin

IX). In the normal resting state of the enzymes, the iron in the haem-iron complex is bound as Fe^{3+} via a thiolate (S^-) from a cysteine residue (Figure 1.12).

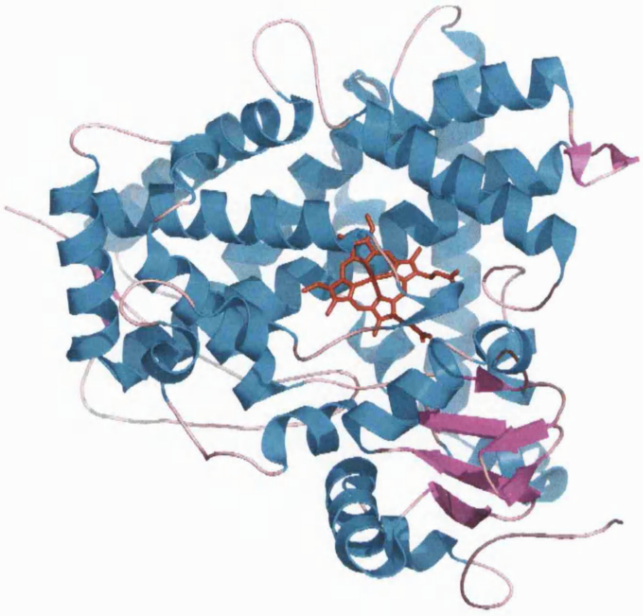


Figure 1.11 Structure of CYP1A2 with the embedded prosthetic haem group. (Made from PDB file 2HI4 with MacPyMol).

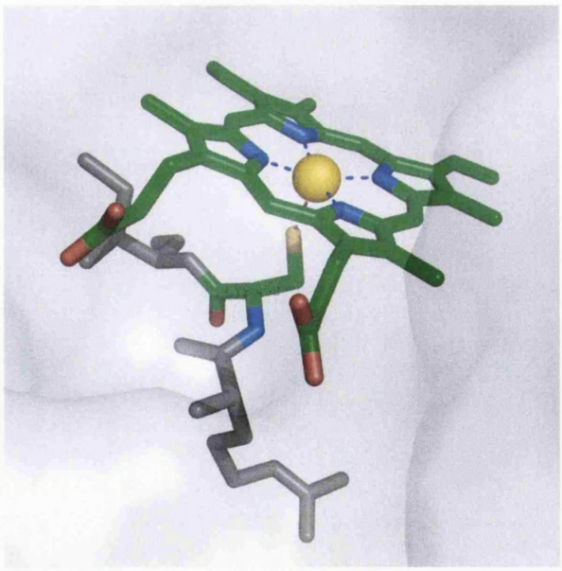


Figure 1.12: Human CYP1A2 haem group: The iron in the centre of the haem group of CYPs forms complexes with the poryphorin groups and in the deprotonated form ligates a S^- from the cysteine in the apoprotein. (Made from PDB file 2HI4 with MacPyMol).

CYPs are considered as monooxygenases, as they transfer one oxygen only to their substrates. The typical reaction can be simplified in the following equation: $SH + O_2 + 2e^- + 2H^+ \rightarrow SOH + H_2O$. The substrate SH is oxidized with the help of a CYP enzyme to become a hydroxyl metabolite. The molecular oxygen is split and one oxygen atom is incorporated into the target and the other one is released as water. In order to catalyse this reaction, CYP enzymes need two electrons as reducing equivalents. Therefore another enzyme is employed. The preferred electron donors are NADPH cytochrome P450 reductases, flavoproteins, containing both FAD and FMN and NADPH, and NADPH together with cytochrome b5 [196, 197]. Two flanking prosthetic groups in NADPH cytochrome P450 reductase assist in the direct transfer of electrons to the haem iron in CYP [196].

The catalytic cycle of CYP-mediated reactions (Figure 1.13) shows the binding of the substrate to the active site of the enzyme. The cytochrome P450 remains inactive until it is bound to the substrate molecule, which then eases the transfer of one electron via NADPH P450 reductase to reduce Fe^{3+} to Fe^{2+}. The substrate-Fe^{2+} complex then binds molecular oxygen (O_2,): oxygen now binds as the sixth ligand instead of the water molecule in the normal state. Once again one electron is transferred to the complex via NADPH P450 reductase, which results in the formation of $Fe^{2+}O_2^-$. This state is most probably in equilibrium with $Fe^{3+}O_2^{2-}$. Subsequently two protons are taken up and the iron-oxygen complex can dissociate in Fe^{3+} and hydrogen peroxide. The exact process for the insertion of the oxygen atom in the substrate is still subject of discussions and not fully understood [198]. Different mechanisms have been suggested [198] how the oxygen is introduced into the substrate. One proposed mechanism is the involvement of a perferryl-oxygen (FeO^{3+}) complex into the mono-oxygen transfer [198]. Research in the oxidising species in CYPs is proposing ferric peroxy-anion (Fe^{III}-OO^-), a ferric hydroperoxy complex (Fe^{III}-OOH) as well as an iron oxene species [199].

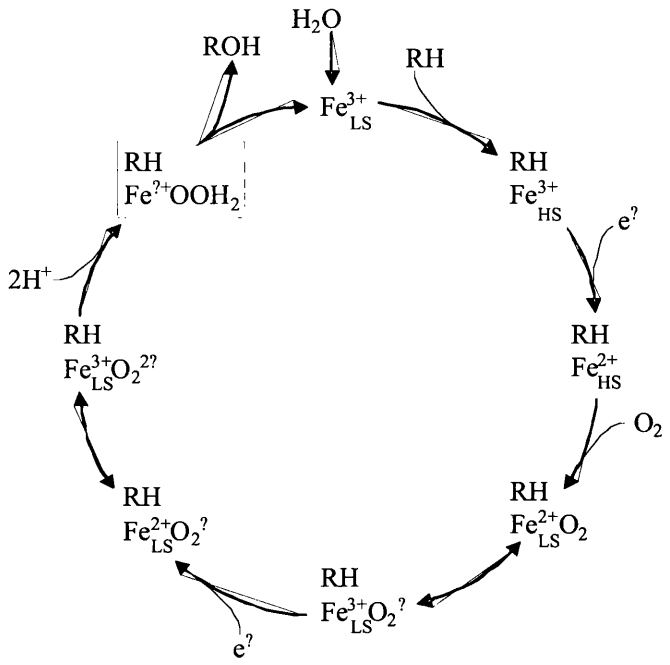


Figure 1.13: Catalytic cycle of CYP enzymes (from Lane [200]). Oxygenation can be divided into different stages. Fe^{3+} binds in a low spin (LS) state to the substrate (RH). Iron passes through the cycle, where it is turned into high spin state (HS). Electrons are transferred via NADPH cytochrome P450 reductase. ROH=product

1.5.2.3 Drug metabolism and pharmacogenetics of CYPs

CYPs of the families of CYP1, CYP2 and CYP3 are responsible for most of the phase I metabolism of xenobiotics and estimations on the numbers of their substrates exceed 200,000. Enzymes of the families CYP4 are mainly involved in fatty acid metabolism and metabolism of some xenobiotics. Families 5 to 51 metabolise endogenous compounds e.g. fatty acids or cholesterol [187]. Of the families 1 to 3, CYP3A4 is the most abundant CYP enzyme present in human liver (Table 1.1) and also contributes to the metabolism of many drugs currently on the market (Figure 1.14, [186], study from 2003). CYPs are able to metabolise a variety of substrates, but it also has to be emphasised that in turn the same or other substrates can induce their activity (Table 1.1). An induction or inhibition of metabolic activity leads to faster or slower metabolism respectively. This can decide the effect of drug therapy and for example in the case of inhibition of a CYP responsible for the metabolism of a drug with a small therapeutic index, can result in high and often toxic blood concentrations of the drug [201]. The extent of metabolic

activity is not only modifiable by drugs, but also by diet, environment, gender or age [202]; it can also be different in individuals due to genetic polymorphism, which are defined as a difference in DNA sequence found in 1% or higher in a population. Roche have marketed the so-called AmpliChip®, a microarray based pharmacogenomic test, which can identify polymorphisms in CYP2D6 and CYP2C19. In general it is being used to distinguish between poor metabolisers (PMs), who lack the functional enzymes; intermediary metabolisers (IMs), who are either heterozygous for one deficient allele or carry two alleles that cause reduced activity; extensive metabolisers (EMs), who have two normal alleles; and ultra-rapid metabolisers (UMs), who have multiple gene copies. The consequences of these polymorphisms signify the physiological importance of CYPs. CYP26A1 for example is involved in the metabolism of retinoic acid, which plays an important role in embryonic development. Abnormalities due to excess or deficiency in retinoic acid are correlated to the expression of the CYP26A1 gene in embryos [203]. Also defective aldosterone biosynthesis is supposed to be connected with CYPs. CYP21 and CYP11B2 are involved in cortisol and aldosterone biosynthesis and thus have an effect on it. Furthermore CYP1B1 can have an influence in formation of primary congenital glaucoma [204]. Consequences of CYP polymorphism, drug induced inhibition and induction etc. can be adverse drug effects, which are estimated to be the cause of over 5,000 deaths in the UK (in 2004 [202]).

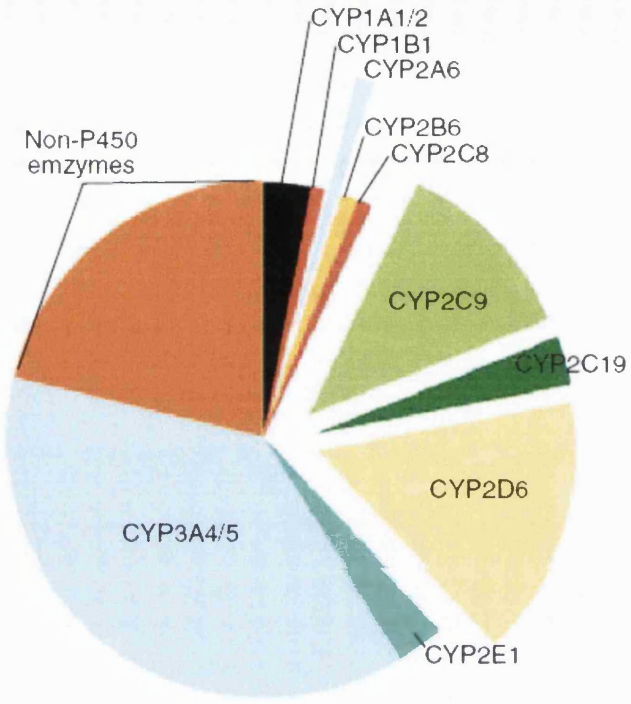


Figure 1.14: Proportion of CYP enzymes involved in metabolism of marketed drugs in 2003. From Guengerich, 2003 [205]. Non-P450 enzymes include alcohol and aldehyde dehydrogenases and flavin-containing monooxygenase etc.

One widely investigated influence of CYPs is their role in cancer development as well as cancer treatment. CYPs are responsible for the bioactivation of pro-carcinogens and other genotoxic xenobiotics. During oxidative metabolism, electrophilic metabolites are formed, which can form adducts to DNA and thus leads to genotxicity [206]. For example, procarcinogens such as azoxymethane or N-nitrosodimethylamine are transformed into active compounds by CYP2E1 which can form DNA adducts [207-209]. The susceptibility can also depend on polymorphism, e.g. an increased risk of lung cancer has been suggested for CYP1A1 MspI polymorphism [168]. Tumour tissues also contain CYPs, which have become targets for anti-caner therapy. There are different strategies to attack CYPs in tumour. A CYP19 (aromatase)-inhibitor is used in the case of breast cancer (Formestane). Also the inhibition of CYP24 in order to keep it from metabolising calcitriol (vitamin 1,25-D_3), which can be used as an anticancer agent, is promising [210]. Another approach is the use of CYPs known to be expressed in a particular tumour to metabolise an inactive pro-drug into the active form selectively in

the tumour. There are already several established anti cancer-drugs, which require CYP-induced activation. Cyclophosphamide and Ifosphamide for example, are used in the treatment of a variety of cancers including testicular, breast, lung, cervical, ovarian, bone cancer as well as lymphoma and some forms of leukaemia. Their induction is achieved by CYP2B6, 3A4 and 2C9. CYP2A6, CYP2C9 and CYP1A1 are involved in the activation of Tegafur, a 5-fluorouracil derivate, which is applied primarily in bowel cancer therapy. Further chemotherapeutics activated by CYP are dacarbazine and procarbazine metabolised by CYP1A1, CYP1A2, CYP2E1 and CYP2B respectively and are used against melanoma, Hodgkins lymphoma and brain tumours. The anti-oestrogen drug, Tamoxifen, is activated by CYP2C9, CYP2D6 and CYP3A4. Taking this into consideration, a future goal in anti-cancer drug development is the further analysis of different kinds of tumours regarding their CYP composition, bearing in mind possible polymorphisms [175].

Table 1.2: Individual CYPs in different tumour types and their expression, taken from Patterson and Murray [175]. Protein expression was primarily determined using immunohistochemistry. n.d.=not determined

Tumour type, number of tumours	CYP expression (% of positive tumours)			
	CYP1A	CYP1B1	CYP2C	CYP3A
Bladder (invasive transitional cell carcinoma), n=25	68%	100% (n=8)	28%	68%
Breast (invasive ductal carcinoma), n=54	40%	80% (n=60)	0%	30%
Colon (invasive adenocarcinoma), n=28	75%	87% (n=80)	n.d.	61%
Ovary (including serous and mucinous carcinoma), n=170	n.d.	80%	n.d.	60%
Secondary ovarian carcinoma, n=48	n.d.	80%	n.d.	n.d.
Kidney (clear cell carcinoma), n=30	100%	100%	0%	100%
Oesophagus, n=50	64%	100%	26%	72%
Prostate (adenocarcinoma), n=51	63%	90% (n=10)	25%	60%
Sarcoma, n=37	70%	90% (n=9)	n.d.	78%
Stomach (adenocarcinoma), n=39	60%	90% (n=10)	0%	30%

1.5.2.4 Human CYP1A2 and CYP2E1

To date CYP enzymes have been studied using a variety of different methods including immunoblotting, PCR and enzymatic activity assays. The studies performed have not provided definitive data on the protein level for individual CYP-isoforms for a number of different reasons. The application of RT-PCR to quantify e.g. CYP2E1 mRNA was

successful [211], yet the results represent no 'real' protein data but only mRNA content. Western blot analysis can be performed at the protein level [212], but is dependent on the availability of specific antibodies, and may not provide 100% selectivity, additionally only one CYP isoform can be quantified per analysis. Studying the enzymatic activity is another approach in CYP studies and increased activities are used as an indicator of higher protein levels. This is used e. g. to measure protein levels of CYP2E1 [213] by studying its enzymatic activity via the metabolism of chlorzoxazon and CYP1A2 by observing the metabolism rate of caffeine. However, there remains the problem of choosing a suitable substrate, as more than one CYP isoform can have a similar enzymatic activity. A further analytical technique, LC-MS/MS, is gaining in popularity for qualitative and quantitative analysis of proteins [214] and in this work we will focus on the development and validation of suitable methods. For the analysis of any CYP it has to be remembered that they are membrane proteins, which makes them hydrophobic and requires special attention. Thus, the established 1D-LC-MS/MS approach for identification and quantification experiments was chosen.

It is obvious that induction or inhibition of these enzymes will result in the change of metabolism of a variety of structurally different substrates. Therefore, the intake of any drug has to be considered when looking for explanations of adverse drug reactions, as the metabolism of even 'simple' substrates such as acetaminophen (paracetamol) or oral contraceptives (17β-ethynylestradiol) can be inhibited by ketoconazol or induced by barbiturates [215].

CYP2E1 only makes up ~7–9% of liver CYP enzymes [216, 217]; yet, it is no less important than higher abundant CYPs as it takes part in the metabolism of a large variety of substrates including aromatic compounds (e. g. acetaminophen, toluene, benzene, chlorzoxazone), alcohol (ethanol), halogenated anaesthetics (e.g. halothane), heterocyclic compounds (e.g. isoniazid, pyridine), some nitrosamines (e.g. N,N-dimethylnitrosamine) [208] and also endogenous compounds such as acetone and fatty acids (linoleic and arachidonic acids) [218]. In turn expression of CYP2E1 can be induced by many of these substances including alcohol. Ethanol is oxidised by CYP2E1 to reactive acetaldehyde and further to acetate [219], and there is a relationship between CYP2E1 and alcoholic liver disease [220]. Although the reasons for alcohol induced liver injury are not completely understood, damage caused by oxidative stress and lipid

peroxidation are considered to be key factors and the impact of CYP2E1 on the accumulation of reactive oxygen species (ROS) has been reported [218, 221]. The conversion of acetaminophen to reactive and toxic electrophilic metabolites is known to have fatal effects. These reactions are catalysed by CYP2E1 together with 1A2 and 3A4 [222]. CYP1A2, representing 15% of total liver CYPs, is known for its influence in the activation of carcinogens in tobacco smoke, into genotoxic compounds. Arylamines and nitrosamines, prominent in tobacco smoke have been associated with CYP1A2 mediated oxidation to toxic products [223]. Furthermore, polynuclear aromatic hydrocarbons (PAH) are reactivated to toxic metabolites via CYP1A2. In addition to the various drug substrates as e.g. lidocain, olanzepine, tacrine or theophyllin, CYP1A2 also metabolises caffeine and melatonin[189].

The widespread involvement of these enzymes as drug metabolisers and the often toxic products that can be formed through them show, that it is important to investigate their distribution further to be able to know where and how much of them is present in different tissues. This information can then be used further to understand and predict drug interactions and the design of tumour targeted drug therapy.

1.6 Human liver proteome

Liver is the main site of metabolism in the human body and contains relative to other organs most CYPs. Human liver microsomes are thus an appropriate matrix to study metabolism. As mentioned above, CYPs can have various influences on metabolism as well as disease development (1.5.2.3). Therefore their investigation is crucial in order to determine the metabolic profile of the liver. The major CYP proteins quantitatively expressed in liver are 1A2, 2A6, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1 and 3A4 [186, 217]. These can be identified in human liver microsomes using a proteomics approach applying 1D-gel coupled to LC-MS/MS [200]

The methods developed in the current work followed the basis of proteomics research, so common procedures as gel electrophoresis, liquid chromatography and mass spectrometry were used. The proteomics analysis on liver has so far been useful to study liver functions and diseases [224]. This way proteomics can assist the discovery of liver malfunctions and aid potentially in the understanding of metabolism and its correlation to disease development. For example in the case of hepatocellular

carcinoma, a kind of cancer which is often diagnosed in a late stage of the disease, proteomics is used to understand the pathogenesis as well as find new biomarkers for earlier detection [225]. Diamond *et al.* recently applied proteomics techniques for the investigation of hepatitis C induced liver fibrosis and found 210 proteins with statistically significant difference associated with the fibrosis state [226]. In a different analysis using proteomics techiques, proteins unique in cases – but not controls – of primary hepatolithiasis were identified in liver tissue [227]. Changes in the proteome of affected tissue are the best source for the development of diagnostic biomarkers, however, samples can only be obtained using invasive procedures and are therefore difficult to use in large scale screening processes. Using changes in the blood proteome as biomarker of disease or risk would therefore be beneficial, however, the large dynamic range of proteins (40 mg/ml for albumin, 5 pg/ml for cytokines) makes the discovery of low-abundant proteins very difficult and requires additional separation and enrichment techniques [228].

1.7 Human lung proteome

Various proteomic studies of different parts of the lung have been undertaken. The studies mostly represent the study of a special part of the lung or special preparation. The analysis of proteins in the bronchoalveolar lavage fluid (BALF) carried out by Noel-Georis *et al* [229] shows the complexity of proteins in the BALF, which presents components originating from the epithelial lining fluid (ELF) that covers the airways. With 2D electrophoresis mapping 93 proteins involved in lipid metabolism, immunological response, tissue repair, cell proliferation and in protecting the lung from microorganisms and oxidative damage were identified. Ostrowski *et al.* [230] undertook a proteomic analysis of human cilia, resulting in the identification of over 200 proteins. The study of lung cancer has become a target for proteomic research, as it is one of the most common human cancers in the western world (Office for National statistics: 30,997 new cases in 2005 in the UK, second most common after breast cancer). There has been a study on human lung squamous carcinoma as this is the most common type of lung cancer [231], which identified forty-one protein spots including several proteins (cyclin D2, XEDAR HSP70 interacting protein, p53-binding protein Mdm2, LIM, cysteine–rich domain protein 1) with up regulation or expression only in squamous carcinoma cells.

Once a cancer marker has been identified, it has to be investigated whether it is unique for one type of cancer, as it is possible for one protein that occurs to be also appearing in other tumours or non-cancer related diseases. This is essential for any kind of biomarker research [232].

Lung is the target organ for every inhaled compound, so concentrations of substances, such as airborne pollutants, tobacco smoke or inhaled medicines are higher in the lung than in the liver and therefore the metabolic capability of lung is of interest. Also bloodborne compounds, which already have been metabolised by the liver, pass through the lung and can be metabolised further; due to the large vascular surface of pulmonary veins, circulating compounds are exposed efficiently to the lung enzymes [233]. Research has been undertaken to identify CYPs expressed in human lung and most of this work has investigated CYPs on the mRNA level or on the protein level by immunoblotting or activity studies [234]. The total amount of CYP in lungs was determined to be about 10 pmol/mg microsomal protein, which is far less than the average amount in liver microsomes (about 300–400 pmol of total CYP/mg of microsomal protein) [235]. Also the localisation of CYPs in the lung is not well known, as over 40 different cell types in the lung are supposed to contain different amounts or types of CYPs. CYP1A1 is primarily expressed in the epithelium of the peripheral airways, and is dependent on smoking status. CYP1B1 has only been identified in alveolar macrophages. In contrast, CYP3A4 seems to be expressed more widely; it has been detected in bronchial, bronchiolar, alveolar epithelium and alveolar macrophages (Table 1.3). The best studied CYP in human lung is currently CYP1A1, but others identified include CYP1A2, 1B1, 2A6, 2C, 2D6, 2E1, 2F1, 2J2, 3A4, 3A5, 3A7 and 4B1. Analysis on the protein level however has still to be thoroughly exploited, and the mRNA levels do not necessarily present the protein level. In addition, the varying place of expression and the influence of smoking status, and maybe also the status of exposure to other pollutants, shows that more detailed, CYP targeted proteomics methods have to be developed.

Table 1.3: Overview of Cytochrome P450s expressed in human lung from Hukkanen *et al.* [236]. Results acquired on the mRNA level are differentiated from results acquired on the protein level. --: moderate negative evidence; -: weak negative evidence; +/-: conflicting evidence; +: weak positive evidence; ++: moderate positive evidence; +++: strong positive evidence

CYP	mRNA	Protein	Comments
1A1	+++	+++	Smokers
	+/-	--	Nonsmokers
1A2	+/-	+/-	
1B1	++	+/-	Protein in alveolar macrophages
2A6	++	+/-	
2B6/7	+++	+++	
2C	+/-	+/-	Protein in serous cells of bronchial glands
2D6	+/-	+/-	
2E1	+++	+++	
2F1	+++		
3A4	+/-	+	Protein in ~20 % of cases
3A5	+++	++	
3A7	+/-		
4B1	+++		

1.8 Conceptual formulation

Metabolising enzymes are of crucial importance for the understanding of disease and development of new treatments and diagnostics methods. To investigate metabolising enzymes on a proteomics basis, several qualitative and quantitative methods are available. The main group of enzymes involved in most of the phase I metabolism in human liver are the cytochrome P450s. They have the ability to modify a variety of endogenous as well as exogenous substrates to enable their excretion, but can also create active or even toxic compounds. The fate of every substance ingested depends on the extent and form of xenobiotic metabolism by specific enzymes.

The aim of this study was to develop, validate and compare qualitative and quantitative methods to investigate xenobiotic metabolising enzymes in human lung and liver.

In particular the following objectives were pursued:

1) Several approaches to overcome the problem of the large dynamic range of protein expression in biological samples were investigated.
2) Methods for the quantification of CYPs in human tissue were developed and validated. To do this, suitable internal standard peptides were chosen in the design of a protocol for absolute quantification.

3) As a proof of concept study, the developed methods were applied to liver and lung tissue and the results are compared to literature data.

This study can deliver reliable methods for the quantification and identification of metabolising enzymes in human tissue, and thereby allow a better understanding of physiological relationships of diseases and potentially find new targets for their therapy and diagnosis.

Chapter 2 Experimental

2.1 Materials

2.1.1 Chemicals

Recombinant Cytochrome P450s (CYPs) 1A2, 2E1 and 3A4 were purchased from PanVera (Invitrogen, Paisley, UK) (purity specification for 2E1 and 3A4 > 85%, for 1A2 > 90%). The heavy isotope labelled peptide GTVVVPTL*DSVLYDNQEFPDPEK (where L* is leucine containing six ^{13}C and one ^{15}N) was synthesised by Thermo Electron Corporation, Hemel Hempstead, UK. The supplier specified the purity to be 95%. The peptides with the amino acid sequence IGSTPVL*VLSR (CYP1A2) and EVTNFL*R (CYP3A4) were synthesised by Sigma Aldrich Company Ltd., Dorset, UK and purity declared to be 99% and 97% respectively.

TEAB (triethylammoniumbicarbonate), HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), sucrose, sodium chloride, iodoacetamide, ammoniumpersulphate (APS), tetramethylethylenediamin (TEMED), Glycerol, bromophenol blue, [Glu^1]-Fibrinopeptide B human, Coomassie brilliant Blue R and G were purchased from Sigma Chemicals, St. Louis, USA. SDS (sodiumdodecylsulphate), Tris-HCl (Tris-(hydroxymethyl)-aminomethane) and DTT (dithiotreitol) were obtained from Melford Laboratories Ltd, Ipswich, UK. Acetic acid, formic acid, phosphoric acid and glycerol were from BDH Laboratory Supplies, Poole, UK. Ammonium bicarbonate and all solvents in HPLC grade were obtained from Fisher Scientific Ltd., Loughborough, UK. The gel electrophoresis reagent ProtoGel (a mixture of 30% *w/v* acrylamide, 0.8% *w/v* N,N'-methylene bisacrylamide, in a ratio of 37.5:1) was bought from National Diagnostics (Hessle Hull, UK). Sequencing grade porcine trypsin was obtained from Promega, Southampton, UK.

The prestained protein marker (6–175 kDa) for self cast SDS-PAGE were purchased as a mixture from New England Biolabs (Hitchin, Hertfordshire, UK), and

contained *E. coli* maltose-binding protein (MBP), β-galactosidase (175 kDa), *E. coli* MBP-paramyosin (83 kDa), bovine liver glutamic dehydrogenase (62 kDa), rabbit muscle aldolase (47.5 kDa), rabbit muscle triosephosphate isomerase (32.5 kDa), bovine milk β-lactoglobulin A (25 kDa), chicken egg white lysozyme (16.5 kDa) and bovine lung aprotinin (6.5 kDa). SeeBlue® Plus2 Pre-Stained standard molecular weight maker used in the precast gels was bought from Invitrogen and consisted of a mixture of Myosin, Phosphorylase B, bovine serum albumin, glutamic dehydrogenase, alcohol dehydrogenase, carbonic anhydrase, myoglobin-Red, lysozym, aprotinin and insulin B-chain.

2.1.2 Subjects

2.1.2.1 Human lung microsomes

Human lung microsomes were processed and provided by the UK Human Tissue Bank (UKHTB, Leicester, UK). Eight samples of human lung tissue-microsomes were supplied in total. All derived from healthy tissue states obtained from post mortems. Samples P1 to P7 are microsomal fractions. Sample P8 derived from the same individual as P7, representing the S9-fraction. This is the fraction containing cytosolic and microsomal material. Detailed information about individual subjects is shown in Table 2.1.

2.1.2.2 Human liver tissue

Human liver tissue samples were provided by the Department of Surgery, Royal Free and University College Medical School, University of London, UK. The protocol for the study was approved by the Ethics Committee of the Royal Free Hospital and University College School of Medicine. Informed patient consent for use of tissues was obtained in all cases. Control sections of liver tissue of liver cancer patients as well as sections of the tumour were collected during surgery and used in this study. Table 2.2 lists the details of the individuals with their medical history. Tissues for the study were accessed from the resected and discarded masses of tumour and surrounding liver, which in the case of subjects 3 and 4 were removed as part of the surgical treatment for hepatic metastases arising from colon cancers. The tumours of subjects 1 and 2 were confirmed to be primary liver tumours. Samples of subjects 6, 7 and 8 were derived

from liver cancer patients as well but no tumour sections were collected. Liver and liver tumour samples were immediately snap frozen in liquid nitrogen and stored at -80°C until further processing.

Table 2.1: Details of human lung samples

Sample	Sex	Age	Smoker	Lung disease
P1	female	60	yes	none
P2	female	20	no	Asthmatic for 5 years
P3	male	55	no	Asthmatic, pollen allergy
P4	female	51	stopped 5.5 years prior death	none
P5	male	29	no	none
P6	female	45	no	none
P7	female	38	no	none
P8	Cytosolic and microsomal fraction from subject 7			

Table 2.2: Detailed information about liver samples; (N = normal liver control, T = tumour)

Sample	Sex	Age	Liver issue type
H1 N	male	41	Liver control
H1 T			Primary liver tumour
H2 N	female	73	Liver control
H2 T			Primary liver tumour
H3 N	female	79	Liver control
H3 T			Colon carcinoma metastases
H4 N	female	60	Liver control
H4 T			Colon carcinoma metastases
H6	female	64	Liver control
H7	male	78	Liver control
H8	male	45	Liver control

2.2 Sample preparation Methods

2.2.1 Microsomal preparation of human liver tissue

Human tissue samples had to be processed according to health and safety rules to avoid direct contact with the tissue. Thus, the preparation of the microsomes was carried out in a Class 2 biological safety cabinet in a HSE Containment Level 2 laboratory. Pieces of liver tissue and (where available) liver tumour were cut, weighed and snap frozen in liquid nitrogen. Sections of approximately 0.5 g were used, as this was a easily manageable amount for processing. The frozen tissue pieces were each ground in a percussion mortar and pestle (manufactured by School of Pharmacy workshop) and the powdered tissue was put in a glass homogeniser to be homogenised with a polytetra-

fluoroethylene-head pestle (Biorad, UK). These steps were performed on ice to prevent activation of enzymes in the sample. For every 100 mg of tissue 1 ml of homogenisation buffer (0.25 M sucrose, 50 mM HEPES, 100 mM sodium chloride) were added. The buffer contained one Complete protease inhibitor tablet (Roche Diagnostics, Welwyn Garden City, UK) per 50 ml. The first centrifugation was performed at 1000 × g for 10 min at 4°C (Sigma 6K10 centrifuge, Sigma laboratory centrifuges, Germany) to sediment nuclei, cell debris and unbroken cell. Subsequent centrifugations were performed in a Beckman Optima L90K centrifuge (Beckman Coulter Inc, USA). The supernatant was then centrifuged at 16,700 × g for 20 min at 4°C. The pellet from this step represents the mitochondrial fraction which was suspended in the storage buffer containing 0.5 M TEAB and 0.1% SDS (pH 8.5 at room temperature) and stored at -80°C. The supernatant resulting from this was centrifuged at 200,000 × g for 70 min at 4°C. The microsomal pellet was suspended in the storage buffer and stored at -80°C.

2.2.2 Bradford assay

The total protein content of the microsomal samples was determined using the Bradford dye (Coommassie Brilliant Blue), which binds to aromatic and basic amino acids [237]. For the protein quantification in human lung microsomes a self-prepared Bradford reagent was used. Therefore 30 mg Coommassie brilliant blue R-250 were dissolved in 100 ml absolute ethanol. After addition of 50 ml concentrated phosphoric acid, the volume was made up to 1 L with ultra pure water (18 MΩ water, ELGA-Purelab Prime, High Wycombe, UK). The solution was filtered to remove any undissolved dye and stored in the dark at 4°C. 100 µl of sample were mixed with 1 ml of this dye and incubated for five minutes. A purchased Bradford reagent "Quickstart Bradford Dye", Biorad, Hemel Hempstead, UK (containing Coommassie brilliant blue G-250) was used to determine total protein content in human liver microsomes, mitochondria as well as in tumour liver microsomes. 1 ml of this reagent had to be mixed with 20 µl of sample and incubated for five minutes. In every assay bovine serum albumin (BSA) (Sigma or Biorad) was used prior to the unknown samples to generate a standard curve (10, 20, 30, 40, 50, and 60 µg/ml). For each sample the absorbance was recorded at 595 nm using a UV spectrometer (Biochrom Ltd, Cambridge, UK). As a blank, 100 µl (or 10 µl in the case of the Biorad dye) ultra pure water was used to set the spectrophotometer to zero.

When necessary, biological samples were diluted to achieve an absorbance value within the standard curve. The protein content was calculated with the measured absorbance and the plotted calibration curve. BSA standards and biological samples were always prepared and measured in pairs.

Recombinant CYP1A2 was diluted with ultra pure water and analysed in the same way to measure total protein content.

2.2.3 SDS-Gel-Electrophoresis of lung microsomes

For the analysis of human lung microsomes "self cast"-gels were used to separate samples on a Hoefer mini gel system (Amersham Bioscience, Little Chalfont, UK). The running buffer consisted of 25 mM Tris, 0.25 M glycine and 0.1% SDS. A discontinuous buffer system was applied by loading the samples onto a stacking gel, containing 5% acrylamide, 0.13 M Tris-HCl pH 6.8, 0.1% *w/v* SDS, 0.1% TEMED and 0.1% *w/v* APS. The stacking gel was polymerised on top of the resolving gel (10% acrylamide, 0.4 M Tris-HCl pH 8.8, 0.08% *w/v* SDS, 0.08% TEMED and 0.08% APS).

25 µg total protein as determined by the Bradford assay were used for analysis per lane. In general, 2 µl of each lung microsomal sample and 9 µl of Laemmli [238] buffer (63 mM Tris/HCl pH 6.2, 10% glycerol, 2% SDS, 0.01% bromophenol blue, 100 mM DTT) were mixed and boiled for 5 min at 95°C. The DTT was used in excess to reduce disulfide bonds, forming free sulfhydryl-groups.

To prevent reoxidation of free sulfhydryl-groups, samples were then alkylated with iodoacetamide (alkylation buffer: 63 mM Tris/HCl pH 6.2, 10% glycerol, 2% SDS, 0.01% bromophenol blue, 200 mM iodoacetamide) for 30 min at room temperature in the dark, using two equivalents of iodoacetamide (relative to DTT). 9 µl of alkylation buffer were added to make it up to a final volume of 20 µl. A prestained protein marker (6-175 kDa, New England Biolabs) was boiled at 95°C for 5 min prior to loading onto the gel. 10 µl of the marker was added per lane (one lane of the gel was loaded with marker). The electrophoresis was started with a voltage of 70 V, 20 mA. After the samples had run over the borders of the stacking gel, these parameters were increased to 110 V, 30 mA. The run was stopped when the dye front had reached the lower end of the gel. Gels were stained with a Coomassie Brilliant Blue R 250 stain (0.2% Coomassie Brilliant Blue R-250, 30% methanol, 10% glacial acetic acid in ultra pure water) for one

hour. To remove the background stain, the gels were washed several times with destaining solution (30% methanol, 10% glacial acetic acid in ultra pure water). For two subjects (No 3 and 4) the whole lane was cut out from top to bottom and divided into 36 bands. For subjects 1, 2, 5–8 only the area in the mass range of 45–62 kDa (molecular weight range of CYPs) was cut out into five bands.

2.2.4 SDS-Gel-Electrophoresis of liver microsomes and mitochondria

Ready-made gels were later chosen to decrease contamination and labour time. The NUPAGE® electrophoresis system from Invitrogen (Carlsbad, USA) with a SDS PAGE buffer system was used. The Samples were heated for 10 min at 70°C with NUPAGE® sample buffer (consisting of 106 mM Tris HCl, 141 mM Tris base, 2% Lithium dodecyl sulphate, 10% glycerol, 0.51 mM EDTA, 0.22 mM SERVA® Blue G250 and 0.175 mM Phenol Red at pH 8.5 as a four times concentrated solution) and NUPAGE® reducing agent (consisting of 500 mM DTT in a ten times concentrated solution). Following this, alkylation was performed in the dark for 30 minutes at room temperature by addition of iodoacetamide to obtain a final concentration of 20 mM. Typically, 25 µg protein (corresponding to 5 µl sample) were loaded on NUPAGE® 4–12% Bis-Tris (Bis-(2-hydroxyethyl)-imino-tris-(hydroxymethyl)-methane-HCl) pre-cast 1.0 mm 10 well gels. At least one lane was loaded with prestained molecular weight marker (SeeBlue® Plus2 Pre-Stained Standard,Invitrogen). NUPAGE® MOPS (3-(N-morpholino) propane sulfonic acid) SDS running buffer was used. Power parameters used were 200 V and 115 mA. When the dye front had reached the bottom of the gel the system was stopped. Gels were stained with SimplyBlue™ Safestain (Invitrogen) using a microwave protocol recommended by the producer. The destaining of the gels was performed in water.

The CYP mass region area (approximately 45–62 kDa) was excised and divided into 5 bands. A big band was dedicated to the area around 50 kDa (band 3) and two smaller bands were cut out above as well as below this area to adjust for uneven cutting. As an additional indicator for migration of CYPs, a recombinant CYP protein (treated in the same way as the sample) was run on the same gel.

2.2.5 Electrophoresis of recombinant proteins

Recombinant CYP1A2, 2E1 and 3A4 were reduced and alkylated as described for the liver and lung samples and loaded on the gel. At least one lane was always loaded with the molecular weight marker. Different amounts of recombinant protein were applied on the gel in a range from 0.75 to 15 pmol. Resulting single bands were cut out.

2.2.6 Electrophoresis of liver microsomes spiked with recombinant protein

Different amounts of recombinant CYP2E1 were added to the microsomal liver preparation of subject number one and treated as described above, including reduction and alkylation before electrophoresis. SeeBlue® Plus2 Pre-Stained Standard was also loaded on the gel and the CYP-mass region was excised.

2.2.7 Cutting and destaining of bands

Bands of interest of self cast gels were cut out and each band was placed in an Eppendorf tube and washed several times with 100 µl ultra pure water until neutral (controlled with pH paper). Thereafter, 150 µl destaining solution (acetonitrile/50 mM aqueous ammonium bicarbonate; 1:1 *v/v*) were added. The samples were vortex-mixed for 30 min and this step was repeated until the bands were transparent. The destained samples were then placed on a razor blade and cut with a second blade into very small pieces (< 1 mm^2). A drop of water was added to keep the pieces on the blade. Bands derived from NUPAGE gels were chopped into small pieces directly as described above and destained with 50 mM aqueous ammonium bicarbonate solution in ethanol (1:1 *v/v*). Derived gel pieces from all gels were then washed and dehydrated twice with 100% acetonitrile and dried in a speed vac (Jouan RC10 22, Jouan Nordic, Denmark) for 30 min with no heat employed. The dried gel pieces were then ready for in-gel-digest.

2.2.8 In-gel digest with trypsin

Sequencing grade trypsin was stored in an acidic stock solution (0.2 µg/µl) at -20°C. To activate trypsin, the stock solution was diluted with 50 mM NH_3HCO_3 (ammonium bicarbonate) to achieve a final concentration of 12.5 ng/µl. This was carried out on ice to prevent autolysis of trypsin. Trypsin stock was prepared for all experiments using this

method. Different amounts of trypsin were added onto the gel pieces – either alone or in combination with a heavy isotope labelled peptide – and left on ice to allow absorption of trypsin into the gel pieces. After that an additional volume of ammoniumbicarbonate was added to prevent gel pieces from drying out during digestion and to assure alkaline pH. The samples were then left in an incubator (Economy incubator, Gallenkamp, UK) at 37°C overnight (usually at least 16 hours).

2.2.8.1 In-gel digest of lung microsomes and extraction

To each sample (band-piece) 20 µl of the trypsin solution were added (corresponding to a total trypsin amount of 250 ng per band). After 30 min, 50 µl of 50 mM NH_3HCO_3 were added and the sample then incubated.

The samples were centrifuged for 5 minutes, using a Biofuge fresco Heraeus bench centrifuge (Kendall Laboratory products, Germany) and the supernatant was taken out and collected in a fresh vial. To extract more of the peptides remaining in the gel pieces, 30 µl of 50 mM NH_3HCO_3 was added. Samples were then vortex-mixed and sonicated for 6 min, centrifuged and the supernatants combined with the previous ones. Thereafter 30 µl of CH_3CN was added and samples extracted as described above; all supernatants were combined. A further extraction with 30 µl 5% formic acid instead of CH_3CN followed. In a final step, the samples were dehydrated with 30 µl CH_3CN and extracted once more. The combined supernatants were dried in a speed vac.

2.2.8.2 In-gel digest of liver microsomes and liver tumour microsomes

For the identification of proteins in these samples dried gel pieces derived from band three (largest band in the P450 mass region) needed more trypsin. Thus, to this one 375 ng of trypsin were added, and 250 ng to the bands 1, 2, 4 and 5.

For the quantification experiments, trypsin was mixed with the required heavy isotope labelled peptide prior to adding this mixture onto the gel pieces. After 30 min on ice, samples were covered with a sufficient amount of 50 mM ammonium bicarbonate and incubated over night.

Extraction was performed by alternating washings with 5% formic acid in 50% acetonitrile and 100% acetonitrile. All extracts were combined and dried completely in a speed vac (no heat employed).

2.2.8.3 In-gel digest of liver microsomes spiked with recombinant protein

To all gel pieces resulting from this experiment a mixture containing the same amount of labelled peptide (4 pmol) and 375 ng of trypsin was added and digested as explained for the samples above.

2.2.8.4 In-gel digest of recombinant protein

All gel pieces from this experiment were digested with a mixture of 4 pmol of corresponding labelled peptide and 250 ng of trypsin and extracted after digestion as described above.

2.2.9 In-solution digest of recombinant protein

Digestion of recombinant CYP2E1, 1A2 and 3A4 (10 µg each) was performed in-solution using 0.2 µg of sequencing grade modified trypsin (Promega, Southampton, UK) in 100 mM aqueous ammonium bicarbonate (trypsin to substrate 1:50). The mixture was incubated over night at 37°C and stored in aliquots at -20°C.

2.2.10 Preparation of extracted peptides for the LC-MS

Dried extracts were redissolved before mass spectrometric analysis in a suitable solvent for each sample. Samples were then centrifuged for 5 min at 16,000 × g to avoid any gel-residues blocking the column.

Extracts from lung experiments and all liver experiments including analysis of CYP1A2 and 3A4 were reconstituted in 0.1% formic acid. Liver samples involving CYP2E1 quantification as well as samples containing recombinant CYP2E1 were redisolved in 100 mM ammonium bicarbonate containing 5% acetonitrile and 0.1% formic acid.

2.3 Liquid chromatography–mass spectrometry

2.3.1 Identification of proteins on the QTOF

For this part of the work a Micromass Q-Tof Ultima Global coupled to a Micromass CapLC System (both Waters, Manchester, UK) was used. Two different settings of the mass spectrometer were compared.

2.3.1.1 Liquid chromatography on nano-LC

Reversed phase liquid chromatography (RP-LC) was carried out using a Micromass CapLC System (Waters, Manchester). The samples were picked up with an auto sampler (5 µl injection volume) and loaded at high flow rate (15 µl/min, using the auxiliary pump C) onto a C18 reversed phase trap column (300 µm × 5 mm PepMap™ C18, 5 µm, 100 µ, LC-Packings/Dionex, Camberly, Surrey, UK) for preconcentration. Two LC runs were conducted; in a 52 minute run after 3 minutes loading time a gradient with mobile phases A and B (see Table 2.3) is applied at a flow rate of 200 nl/min. The analytical column was equilibrated with 94% A and 6% B and three minutes after sample loading the ratio of B was increased to 30% over 27 minutes and then to 80% over 5 minutes. For the next 5 minutes B was kept constant at 80% and A at 20%. The proportion of B was decreased over two minutes back to 6% (A at 94%) and maintained for re-equilibration for 10 min (Table 2.4) Thereby peptides are eluted from the trap column onto the analytical column (Waters Symmetry®, 75 µm ID × 150 mm). The sample make-up solvent was 0.1% formic acid. The proportions for the 63 min gradient are shown in Table 2.5 and Figure 2.2.

Table 2.3: Mobile phases used for liquid chromatography.

Mobile phase	H_2O	CH_3CN	Formic acid
A	95%	5%	0.1%
B	5%	95%	0.1%
C	99.9%	0%	0.1%

Table 2.4: Gradient system used for 52 minute run. For mobile phases see Table 2.3

Time [min]	%A	%B	Flow [µl/min]
0.1	94	6	5
3.0	94	6	5
30	70	30	5
35	20	80	5
40	20	80	5
42	94	6	5
52	94	6	5

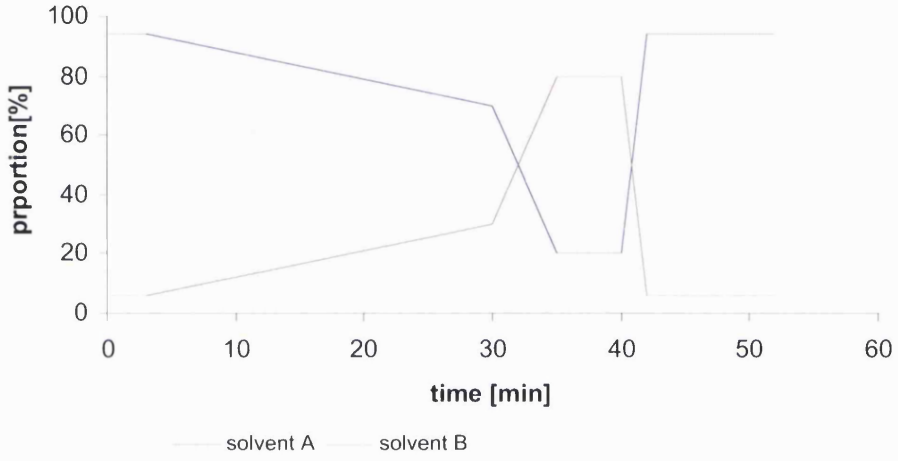


Figure 2.1: LC gradient of 52 min LC-run. The proportion of the hydrophilic solvent A is reduced during the first 35 min and the organic phase is increased for gradually elution of peptides from the column. A washing period with high organic phase continues until 40 minutes run time and is followed by equilibration.

Table 2.5: Gradient system used for 63 minute run. For mobile phases see Table 2.3.

Time [min]	%A	%B	Flow [µl/min]
0.1	94	6	5
3.0	94	6	5
40	72	28	5
45	20	80	5
52	20	80	5
53	94	6	5
63	94	6	5

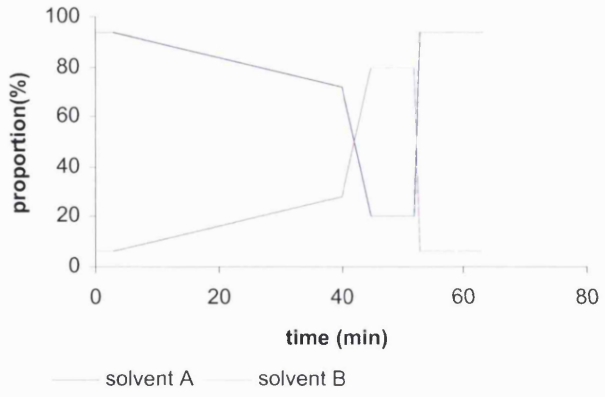


Figure 2.2: LC gradient of 63 min LC-run. The proportion of the hydrophilic solvent A is reduced during the first 45 min and the organic phase is increased for gradually elution of peptides from the column. A washing period with high organic phase continues until 52 minutes run time and is followed by equilibration.

2.3.1.2 Liquid chromatography of recombinant CYP1A2

For the analysis of recombinant CYP1A2 a 36 min LC-gradient was used. The proportion of solvent B was increased from 5 to 35 % over 18 minutes to elute hydrophobic peptides from the column. Solvent B was further increased to 80 % over the next two minutes and the column was washed for 5minutes. A 10 minute equilibration phase was followed prior to the next sample injection.

Table 2.6: Gradient system used for 36 minute run. For mobile phases see Table 2.3.

Time [min]	%A	%B	Flow [µl/min]
0.1	95	5	5
3.0	95	5	5
18	65	35	5
20	20	80	5
25	20	80	5
26	95	5	5
36	95	5	5

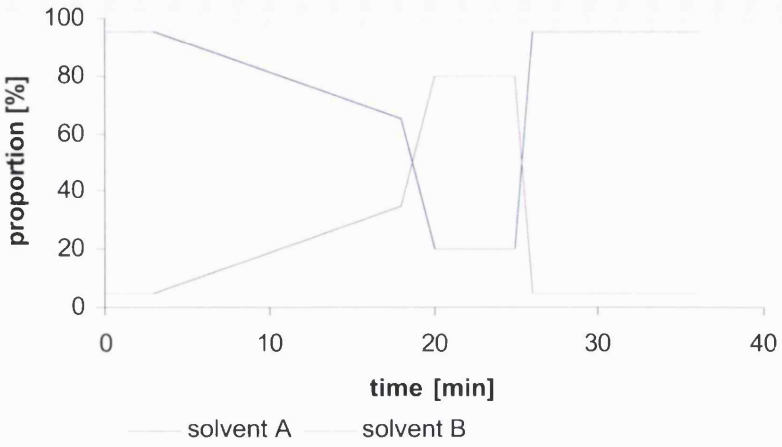


Figure 2.3: LC gradient of 36 min LC-run. The proportion of the hydrophilic solvent A is reduced during the first 20 min and the organic phase is increased for gradually elution of peptides from the column. A washing period with high organic phase continues until 26 minutes run time and is followed by equilibration for 10 minutes.

2.3.1.3 Tandem MS of human lung microsomes:

The capLC system was connected to a Micromass Q-Tof Ultima Global (Waters, Manchester,UK), consisting of a nano electrospray ion source and a quadrupole coupled to a time of flight mass analyser. Ionisation was always carried out in positive ion mode. Parameters used are shown in Table 2.7. For the analysis of the lung samples each digested band sample underwent 4 runs in total, differing in the *m/z* range scanned. Method 1 consisted of a MS/MS run only. Method 2 refers to three consecutive runs with MS/MS 2, MS/MS 3 and MS/MS 4. Thus a sample, which was run with method 1 and 2, was subjected to an overall of 4 runs or 4 injections respectively. The total scan times were 5 to 45 (method 1) or 3 to 60 minutes (method 2). The analysis were performed in data dependent analysis (DDA) mode, starting with a 1 s MS survey scan over the *m/z* range 400–1500 for method 1 and 400–600, 600–800 and 800–1000 for method 2 respectively (listed in Table 2.7). This was followed by a 1 s MS/MS scan with 0.1 s interscan time for up to 4 of the most abundant precursor ions. The intensity threshold for the MS intensity was 10 counts per second (cps). The MS/MS acquisition was carried out for 3.3 s for MS/MS 1 (*m/z* range 50-1800) or 4.4 s

for MS/MS 2–4 (m/z range 50–1700). If the intensity of a MS/MS signal was falling below 4 cps (MS/MS 1) or 3 cps (MS/MS 2–4) the instrument switched back to MS acquisition. The MS/MS 1 method was combined with the 52 min LC-gradient and MS/MS 2, 3 and 4 were each run with the 63 min LC-gradient as described in 2.3.1.1.

Table 2.7: Parameters used for mass spectrometric analysis.

	Run description			
	Method 1	**Method 2**		
Parameter	**45 min MS/MS1**	**60 min MS/MS 2**	**60 min MS/MS 3**	**60 min MS/MS 4**
MS survey scan range [m/z]	400–1500	400–600	600–800	800–1000
MS/MS range [m/z]	50–1800	50–1700	50–1700	50–1700
Total scan time [min]	5–45	3–60	3–60	3–60
MS Survey scan time [s]	1	1	1	1
MS Survey interscan time [s]	0.1	0.1	0.1	0.1
MS intensity threshold [cps]	10	10	10	10
MS/MS scan time [s]	1	1	1 s	1
MS/MS interscan time [s]	0.1	0.1	0.1 s	0.1
Number of components	4	4	4	4
MS/MS to MS switch after [s]	3.3	4.4	4.4	4.4
MS/MS to MS switch criteria	Intensity falling below threshold	Intensity falling below threshold	Intensity falling below threshold	Intensity falling below threshold
Switchback threshold [cps]	4	3	3 cps	3 cps

2.3.1.4 MS/MS of recombinant CYP1A2

In-solution digest of recombinant CYP1A2 containing labelled internal standard peptide specific for CYP1A2 was submitted to 35 minute data dependent MS/MS analysis. Parameters are summarised in Table 2.8.

Table 2.8: MS/MS parameters applied for the MS/MS analysis of in-solution digests of CYP1A2 samples.

Parameter	35 min MS/MS
MS survey scan range	400–1500 m/z
MS/MS range	50–1700 m/z
Total scan time	6–35 min
MS Survey scan time	0.5 s
MS Survey interscan time	0.1 s
MS intensity threshold	10 cps
MS/MS scan time	0.5 s
MS/MS interscan time	0.1 s
Number of components	3
MS/MS to MS switch after	0.8 s
MS/MS to MS switch criteria	Intensity falling below threshold
Switchback threshold	3 cps

2.3.1.5 Calibration of the QTOF

The calibration of the QTOF is very important for the mass accuracy of the instrument and the consistency of the obtained data. GluFib ((Glu^1)-Fibrinopeptide B human, Sigma, USA) was used for calibration. Glufib was directly infused via a syringe and a syringe pump built in to the instrument with a flow rate of 0.3 µl/min. This way the LC was bypassed and only the state of the mass analyser could be observed. GluFib consists of fourteen amino acid residues (amino acid sequence: EGVNDNEEGFFSAR, MW: 1569.6695 Da). First a MS spectrum was acquired to assess the intensity of the mass peak of GluFib (785.842 *m/z*, $[M+2H]^{2+}$); the intensity was checked and intensities of around 400 cps were accepted (maximum 700 cps). This peak was then fragmented by MS/MS and the experimental fragmentation pattern was compared to the theoretically expected one. The instrument was calibrated with the fragment masses in the MS/MS spectrum: *m/z* 72.0813, 175.1195, 187.0719, 246.1566, 333.1887, 480.2570, 627.3254, 684.3469, 813.3895, 942.4321, 1056.4750, 1171.5020 and 1285.5448, all singly charged. Data was only acquired when the calibration was within the instrument specification. Furthermore BSA (bovine serum albumin) was used to monitor the instrument performance taking also into account the chromatographic efficiency. An in-solution digest of BSA was subjected to MS/MS method 1 and the resulting MS/MS data was subjected to database search. The resulting sequence coverage for a 50 fmol BSA injection had to be above 40% for acceptance. BSA was used as a quality control standard to ensure good instrument performance. Actual sample data was only acquired when both tests were passed.

2.3.1.6 Creation of peak-lists:

The resultant MS/MS data was processed into peak list (.pkl) files using the instrument control software Mass Lynx (Waters, Manchester, UK). Spectrum selection criteria were used to combine sequential scans with the same precursor and to process combined scans with a quality assurance threshold of 10.00. The spectra were smoothed (Savitzky-Golay [239] smoothing algorithm, 1×4 channels) and centroided (80%, centroid top; minimum peak width at half height: 4).

2.3.2 Identification and Quantification of CYP enzymes in human liver microsomes

Method development and quantitative work on CYP1A2, 2E1 and 3A4 as well as the identification of CYP1A2 and CYP2E1 in human liver samples were performed on a LCQ^{duo} ion-trap mass spectrometer (Thermo Electron, Hemel Hempstead, UK) coupled to a LC Packings Ultimate Capillary HPLC system with FAMOS™ autosampler (Dionex, Camberley, Surrey, UK).

2.3.2.1 Liquid chromatography on cap-LC

General analysis of the nano-LC was performed by the injection of the sample (1 or 5 µl) via a 20 µl sample loop with 0.1% formic acid as loading buffer assisted by a separate Ultimate Micropump. The sample was loaded on a pre-column (5 mm x 300 µm PepMap™ C18 guard column, 5 µm, 100 µ; Dionex, Camberley, Surrey, UK). After the sample had been washed for 3.5 minutes on the pre-column, it was transferred to the analytical column (PepMap™ C18, 75 µm x 150 mm; 3.5 µm,100 µ) which was equilibrated with 95% solvent A (5% acetonitrile in 0.1% formic acid, Table 2.9) and 5% B (80% acetonitrile in 0.1% formic acid) at a flow rate of 200 nl/min. For the identification and quantification slightly different gradients were used. To avoid cross contamination between the individual sample-injections, blank injections of isopropanol were performed after each sample (three blank injections for 2E1 and two injections for 1A2 analysis) utilising the maximum injection volume (6.4 µl) possible with the sample loop (20 µl) used.

Table 2.9: Mobile phases used on the LCQ

Mobile phase	H_2O	CH_3CN	Formic acid
A	95%	5%	0.1%
B	20%	80%	0.1%
C	99.9%	0%	0.1%

2.3.2.2 LC gradients on LCQ

65 min gradient

For the LC-MS/MS identification of proteins and the standard BSA MS/MS run a 65 min gradient programme was used. After 3.5 min of washing of the trap column, the

proportion of mobile phase B was increased linearly to 28% over 39 min and raised to 80% in the following 5 min. B was kept at 80% for 7 min after which the column was equilibrated for 10 min with 95% solvent A and 5% B prior the next injection Table 2.10, Figure 2.1).

Table 2.10: 65 minute gradient. For mobile phase see Table 2.9

Time [min]	%A	%B	Flow [µl/min]
0	95	5	0.2
3.0	95	5	0.2
42	72	28	0.2
47	20	80	0.2
54	20	80	0.2
55	95	5	0.2
65	95	5	0.2

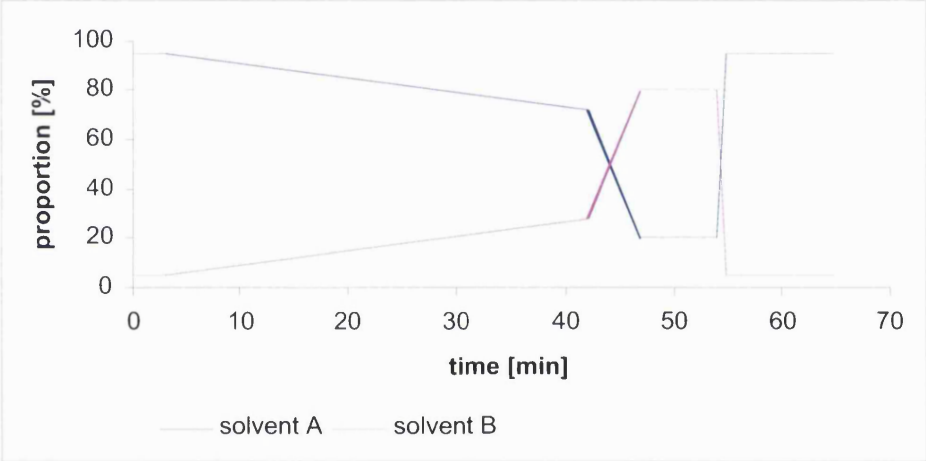


Figure 2.4: LC gradient of 65 min LC-run. The ratio of mobile phase B was increased over 39 minutes in the 65 min gradient to 28% and after a 5-minute rise to 80% this was kept for 7 min to wash the column. From 55 to 65 min the column was equilibrated with 95% aqueous phase A.

58 min gradient:

The seven minutes shorter gradient for the quantification of CYP2E1 implies the linear raising of the proportion of mobile phase B to 40% over 35 minutes and to 80% in 30 seconds to be maintained at this concentration for 9 minutes. The column was re-equilibrated for 10 minutes at 95% A and 5% B (Table 2.11, Figure 2.5).

Table 2.11: 58 minute gradient. For mobile phase see Table 2.9

Time [min]	%A	%B	Flow [µl/min]
0	95	5	0.2
3	95	5	0.2
38	60	40	0.2
38.5	20	80	0.2
47.5	20	80	0.2
48	95	5	0.2
58	95	5	0.2

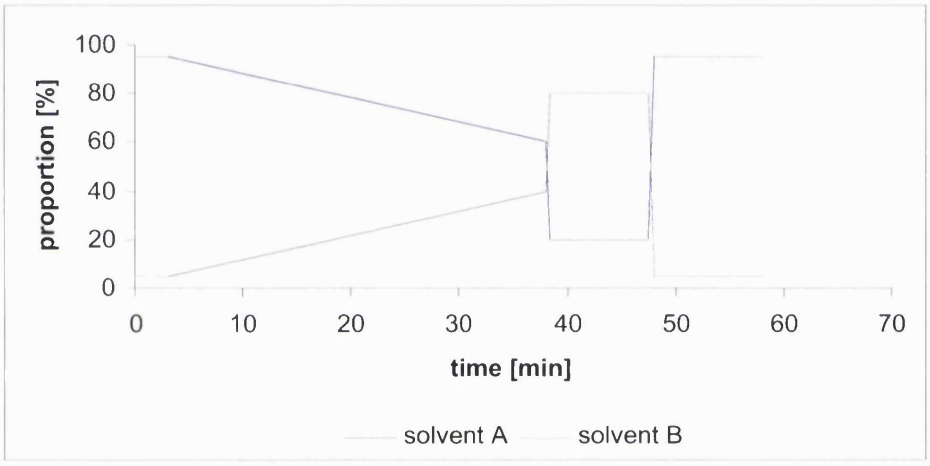


Figure 2.5: LC-gradient of the 58 min LC-run. The proportion of the organic mobile phase A is increased over 35 min to 40% and then kept at 80% for a 10 min wash of the column. After that the column is re-equilibrated for 10 min with high aqueous mobile phase proportion (95% A).

38 min gradient:

This gradient was used for the quantification of CYP1A2 as well as for the blanks and a standard BSA MS run. Here the gradient was 38 minutes with an initial raise of mobile phase B to 40% during 12 minutes and 80% in the next 30 seconds. After 27.5 minutes B is restored to 5% and kept at this ratio for the last 10 minutes of the gradient (Table 2.12, Figure 2.6).

Table 2.12: 38 minute gradient. For mobile phase see Table 2.9

Time [min]	%A	%B	Flow [µl/min]
0	95	5	0.2
3.0	95	5	0.2
15	60	40	0.2
15.5	20	80	0.2
27.5	20	80	0.2
28	94	6	0.2
38	94	6	0.2

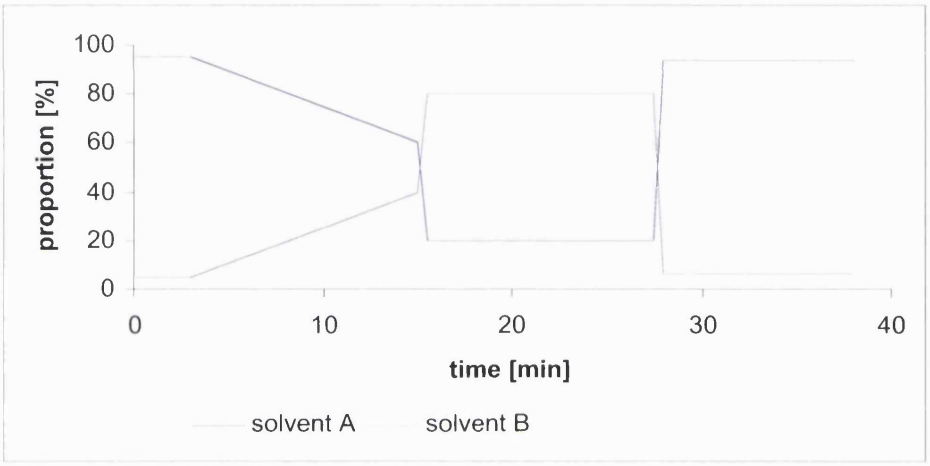


Figure 2.6: LC gradient of 38 min LC-run. The 38 minute gradient on the LCQ consisted of an increase in organic mobile phase over 12 minutes to 40% B and a following 12 minute wash step with high organic mobile phase (80% B). The equilibration of the column before the next run took 10 min.

2.3.2.3 MS analysis on LCQ

The ESI nanospray source was used for all MS analysis on the LCQ^{duo} ion-trap mass spectrometer. ESI was always carried out using the following conditions: positive ionisation mode; spray voltage 1.9 kV; capillary voltage 46 V; capillary temperature 180°C; no sheath or auxiliary gas used. MS analysis was conducted in different modes. All data was collected in centroid mode. All full MS and MS/MS scans consisted of three microscans, each injected for 200 ms at maximum.

The chromatographic performance of the system was monitored each day by comparing retention time and peak shape of a standard. For this purpose a 38 min gradient of an in-solution digest of BSA was analysed. The full scan mode in a range of m/z 400–2000 was used.

Tandem MS

MS/MS analysis was performed in a data dependent mode to identify proteins. Over the 65 minutes run time four scan events were repeated:

1. Full scan, in the *m/z* range 400–2000.
2. Data dependent MS/MS on the most intense ion from 1.
3. Data dependent MS/MS on the second most intense ion from 1.
4. Data dependent MS/MS on the third most intense ion from 1.

The MS/MS collision energy was set to 35%. The minimum signal required for MS/MS data collection was set to 100,000 cps. An ion was put on an exclusion list for 1.25 min (exclusion mass width 2.00) if it appeared as one of the top three most intense ions for more than 3 full scans within 30 seconds. The following contaminant peaks were set on a reject mass list: *m/z* 415.95, 429.95, 445.92, 462.87, 503.92, 519.87, 536.95, and 610.99.

2.3.2.4 Multiple reaction monitoring MS

Quantification of CYP2E1

During the 58 minute run time 3 scan events were repeated:

1. Full scan of *m/z* 700–1400.
2. MS/MS of *m/z* 1281.8, scanning for fragments in the *m/z* range 350–1800
3. MS/MS of *m/z* 1285.3, scanning for fragments in the *m/z* range 350–1800

Quantification of CYP1A2 and 3A4

CYP1A2 was analysed using a shorter run time of 38 minutes. The following scan evens were performed:

1. Full scan of *m/z* 400–2000
2. MS/MS of *m/z* 571.7, scanning for fragments in the *m/z* range 155–2000
3. MS/MS of *m/z* 575.3, scanning for fragments in the *m/z* range 155–2000

For simultaneous quantification of CYP1A2 and CYP3A4, two additional MS/MS scans were performed:

4. MS/MS of *m/z* 878.7, scanning for fragments in the *m/z* range 155–2000
5. MS/MS of *m/z* 885.8, scanning for fragments in the *m/z* range 155–2000

The isolation width of the parent ion in all methods was set to ±1.5 units and the collision energy was 35%. The method was applied to the samples as well as the blanks. For the MRM chromatograms a product ion window of 1 *m/z* unit was selected and peak areas were used in quantitative calculations.

2.3.2.5 Tuning on LCQ

Optimised instrument parameters for successful performance were chosen before sample analysis. Therefore GluFib ((Glu^1)-Fibrinopeptide B) was used as the tuning standard and directly infused via a syringe pump into the nano-ESI source. The fragmentation in MS/MS mode of the parent ion (786.3 *m/z*, $[M+2H]^{2+}$) into its known fragments of *m/z* 72.1, 175.1, 187.1, 246.2, 333.2, 480.3, 627.3, 684.3, 813.4, 942.4, 1056.5, 1171.5 and 1285.5, all singly charged was used. Via the autotune function, instrument parameters were adjusted and the created tune file was used as reference for LC-MS run methods.

For CYP2E1 quantification, optimization of the instrument with a solution of the heavy isotope labelled peptide (GTVVVPTL*DSVLYDNQEFPDPEK) ($[M+2H]^{2+}$, theoretical *m/z* 1285.1, experimental *m/z* 1285.3) was performed. For CYP1A2 and CYP3A4 method development and quantification, the Glufib tune-method was used as for other samples.

2.3.3 Data analysis

2.3.3.1 Database searching

The pkl files derived from the QTOF experiments and raw data obtained from MS/MS analysis on the LCQ (Mascot automatically processes the raw data into .dta files which are merged into a single file suitable for Mascot search), were submitted to the Mascot In-house search engine (Matrix Science Ltd., London, version 2.2); Table 2.13 lists the search parameters. Database searching was restricted to tryptic peptides of human proteins using IPI (International Protein Index, IPI HUMAN, version3) [240]. As variable modifications carbamidomethyl cysteine, methionine oxidation, pyro-glu (N-term) and N-acetylation (protein) were chosen. The precursor and fragment mass tolerance for the QTOF data were 0.3 Da and for the LCQ data precursor mass tolerance 1.0 and for the fragment mass tolerance 0.6 Da were chosen. One missed cleavage was allowed.

Before a protein hit by Mascot was accepted, the Mascot score, giving an estimate for the quality of the identification, and the sequence coverage of the particular protein as well as the peptides and their individual scores were analysed. The score was calculated by the Mascot-System and is based on the absolute probability (P) that the observed match between the experimental data and the sequences in the database is a random event. The score presented in the Mascot results is defined as $-10\log_{10}(P)$, i.e. a low probability for a random event is reflected by a high score. Additionally, a threshold value for the score to evaluate the identity of a protein provided with the search results was used. Proteins below this threshold were checked manually for their peptide information. This way peptide fragmentation and the assignment of the theoretical to the experimental data were rechecked. The recognition of consecutive y-ions and the assignment of peaks of high intensity were another important criteria for a confident identification.

In order to obtain reliable identifications, only peptides giving scores greater than 35 were accepted, if not stated otherwise. Additionally, the false positive rate was determined [68]. Therefore an automatic decoy search was performed against a randomised database with a default significance threshold of $p<0.05$. With precursor and fragment mass tolerance set to the values mentioned above, the false discovery rate at the identity threshold had to be below 5% for acceptance.

For further data analysis and sorting of identified proteins the Proteincenter-software (Proxeon, Odense, Denmark, version 1.2) was used to arrange data after database search. Exported search results from Mascot search were imported into Proteincenter as .csv files. Data was then sorted according to their GO (gene ontology) affiliation into groups of molecular function, biological processes and cellular components.

Table 2.13 Parameters for the Mascot search engine

Data derived from	QTOF	LCQ
Data Input	Masslynx-pkl	Xcalibur-Raw data (.dta)
Type of scarch	MS/MS ion search	MS/MS ion search
Enzyme for cleavage	Trypsin	Trypsin
Variable Modifications	Carbamidomethyl (C), N-Acetyl (Protein), Oxidation (M)	Carbamidomethyl (C), N-Acetyl (Protein), Oxidation (M)
Mass Values	Monoisotopic	average
Protein Mass	Unrestricted	unrestricted
Precursor Ion Mass Tolerance	0.3 Da	1.0 Da
Product Ion Mass Tolerance	0.3 Da	0.6 Da
Max. Missed Cleavages	1	1
Instrument Type	ESI-QUAD-TOF	ESI-TRAP

2.3.4 Uniqueness of peptides

For the decision about uniqueness of one peptide for one specific protein, Blast searches (Basic Local Alignment Search Tool) were performed. Thereby, a given amino acid sequence is aligned to sequences in a database using a set of search algorithms (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). In this work, the blastp algorithm (protein-protein blast) was used to search peptide sequences derived from Mascot-database search against a database of non-redundant (nr) human protein sequences. The results show the alignment between amino acid sequences of the entered sequence and peptide sequences of proteins in the chosen database.

2.3.5 Quantification

For quantification, peaks in the MRM chromatogram for the respective transitions were smoothed using a Gaussian algorithm (number of smoothing points: 7). Peak areas were determined manually using the area label option in Thermo Finnigan Xcalibur QualBrowser software. The ratio of peak areas for these transitions were calculated and used to determine the ratio of native peptide/internal standard peptide.

2.3.6 Statistical analysis

Statistical analysis (*t*-test and ANOVA) were performed using Prism (GraphPad Prism for Mac OS X, GraphPad Software, San Diego, Ca, USA), SPSS 16 (SPSS Inc. Chicago, Il, USA) for Mac OS X and R [241].

Chapter 3 Identification of xenobiotic metabolising enzymes in human lung tissue

3.1 Introduction

The existence of metabolising enzymes in human lung tissue was investigated by analysis of microsomal preparations. Metabolism in the lung is a relevant issue to explore, as the lung is a place of metabolism especially for any inhaled substances, as well as substances in the blood circulation [233]. Various phase I and phase II-metabolising enzymes are involved in lung metabolism [4]. These have to be studied as metabolic toxification or detoxification have an impact on the fate of xenobiotics and this in turn could have an influence on the pathogenesis of diseases [242]. As microsomes represent vesicles derived from the endoplasmic reticulum through subcellular fractionation, they can be expected to contain metabolising enzymes[102]. A 1D-SDS-PAGE protocol [243] in conjunction with tandem mass spectrometry (MS/MS) analysis of in-gel digests was adapted for the study of lung microsomes. With the in-gel digestion method and using a LCQ-Duo ion trap MS/MS instrument it has been previously demonstrated that CYP enzymes, which are one of the most prominent enzyme-families in xenobiotic metabolism, can be identified in human liver microsomes [102].

Here, we use a similar methodology to analyse xenobiotic metabolising enzymes present in lung tissue, however the MS/MS instrument was of the QTOF type rather than an ion-trap. Due to the large differences in concentration of these proteins, data-dependent acquisition methods often result in an underestimation of enzymes present. The aim of this study was to develop a sensitive identification method which overcomes this limitation by dividing the total mass range into smaller sections.

3.2 Microsomal protein content

Eight human lung microsomal samples were prepared and provided by the UK Human Tissue Bank (UKHTB, Leicester), sample 8 is the S9 fraction of sample 7, thus consisting of microsomal and cytosolic fractions. Prior to analysis the total protein content of the eight individual samples supplied was determined using the Bradford method. Due to individual variation of concentration of the microsomes, the samples had to be diluted prior to the assay. Usually dilutions of 1:50, 1:100, 1:500, and 1:1000 were tested to achieve absorbance values within the standard curve. BSA standard curve measurements were carried out in duplicate and absorbance plotted against the concentration (R^2= 0.974, Figure 3.1).

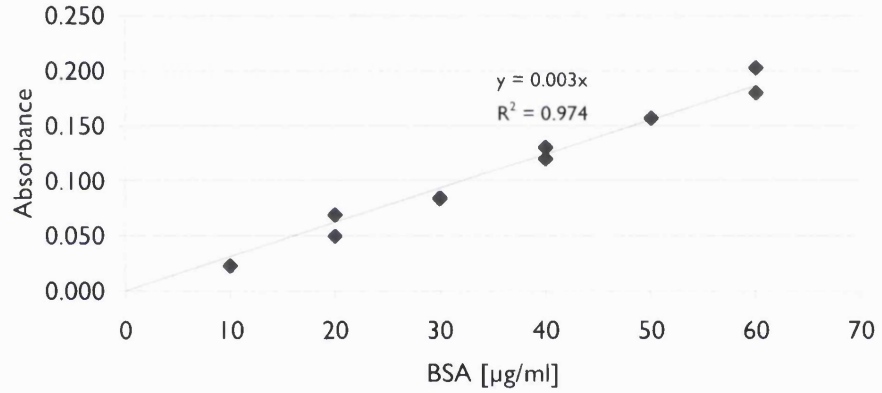


Figure 3.1: BSA standard curve for Bradford assay, measured at 595 nm; BSA concentration in µg/ml

The microsomal protein concentrations shown in Table 3.1 only present the total content of all proteins detectable by Coommassie G-250 in each sample, requiring further methods for identification and/or quantification of any particular protein.

Table 3.1: Protein concentration [µg/µl] of eight human lung microsomal preparations. The protein concentrations are always calculated as mean of duplicate measurements.

Sample	Protein [µg/µl]
P1	4.3
P2	7.4
P3	17.2
P4	3.1
P5	16.0
P6	5.2
P7	14.61
P8	23.5

3.3 Electrophoresis of human lung proteins

The eight samples were each reduced with DTT and alkylated using iodoacetamide as described in materials and methods and separated using self made SDS-PAGE. The total protein amount loaded per lane ranged between 20 and 30 µg. Table 3.2 gives an overview of the amount of protein used for gel electrophoresis.

Table 3.2: Amount of microsomal protein loaded on the gel; sample eight is the S9 fraction of sample seven

Sample	Protein amount loaded on gel [µg/lane]
P1	26
P2	29
P3	30
P4	20
P5	30
P6	26
P7	29
P8	24

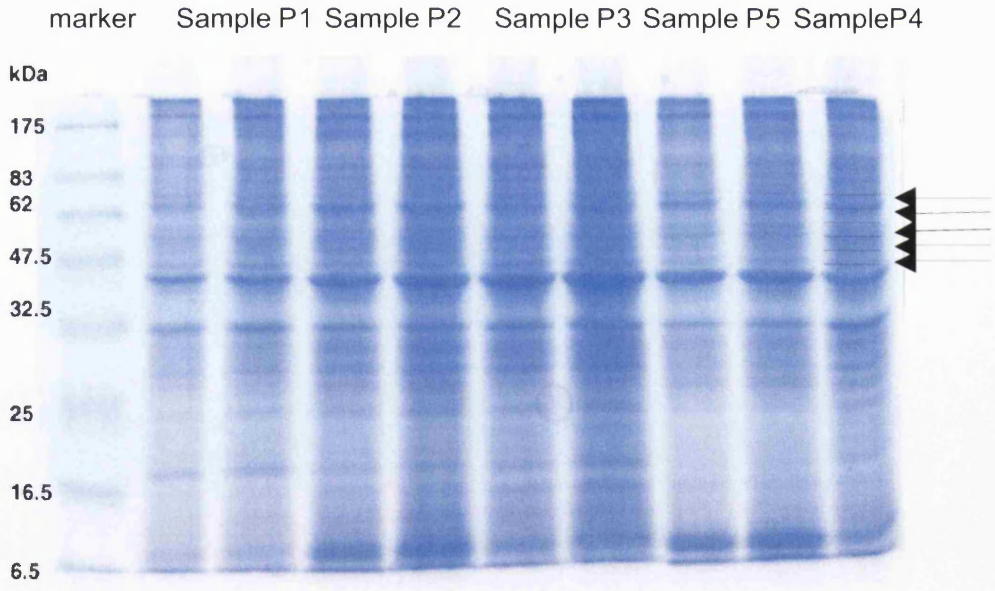


Figure 3.2: SDS-PAGE of lung microsomes (samples 1-5). Each sample was applied in two different concentrations, the right lane at the concentration as shown in Table 3.2; for the left lane, half the sample amount was used. Arrows on the right show the CYP cut-out-area. For identification experiments the higher concentrated lanes of each sample were cut out.

Figure 3.2 shows a SDS-PAGE of five different samples and the arrows on the right illustrate how the bands were cut out. The mass range of these five bands corresponded to the molecular weight range of Cytochrome P450 monooxygenases (CYPs), as the main CYPs have molecular weights in this range. CYPs are one of the most important groups of metabolising enzymes, and are therefore used as a starting point for investigation.

3.4 MS/MS analysis of lung proteins

As indicated in Figure 3.2 by the arrows on the right side, the area, where CYPs were expected due to their molecular weight, was cut out, de-stained and submitted to in-gel tryptic digestion overnight. This was done for all eight samples and the extracted peptides were then analysed by data dependent MS/MS, as described in materials and methods. The experimental data was searched against the IPI human database using the MASCOT search engine. The data of the eight samples was all processed in the same way and peaklists of all eight samples were combined and submitted to database search. The merged data is presented in appendix A (table i). In total 212 proteins were identified, of which 85 were identified by a single peptide. Just one CYP enzyme, namely CYP5A1, equivalent to thromboxane A synthase 1, was identified in patients P1 and P2. One peptide, SVADSVLFLR (peptide score 58) was assigned to CYP5A1 and determined by a BLAST search to be a unique peptide for this protein. However, database search of combined raw data of bands in the CYP450 molecular weight area resulted in the identification of microsomal proteins such as protein disulfide isomerase (IPI00064193), calreticulin (IPI00383751), liver carboxylesterase (IPI00010180) or epoxide hydrolase 1 (IPI00009896) and CYP-related NADPH-cytochrome P450 reductase (IPI00470467) were identified. Lung specific proteins like pulmonary surfactant-associated protein A1 (IPI00012889) or acylglycerophosphocholine O-acyltransferase 1 (Lung-type acyl-CoA:lysophosphatidylcholine acyltransferase 1, IPI00171626) were discovered as well.

3.5 Control of method

The first experiment on lung microsomes identified just CYP5A1 and no other CYPs, although from the literature it was expected that other CYPs were to be found in lung

tissue [162]. Therefore it had to be investigated whether these results were due to instrumental differences, as the experiment was carried out on the QTOF, – a different instrument than that used for the liver experiments conducted previously in our group [243]. In order to investigate whether the in-gel digestion method coupled to the tandem-MS analysis on the QTOF was capable of identifying CYPs, two control experiments were conducted. Recombinant CYP2E1, a CYP present in liver and lung [244], as well as one liver microsomal sample were both analysed to assess the sensitivity and reproducibility of the method.

3.5.1 Validation with recombinant CYP

As a validation sample, recombinant protein was treated in the same way as a lung sample and its data processed likewise. In addition to acting as a positive control, the recombinant protein could provide information on the sensitivity of the assay developed. Recombinant CYP2E1 was chosen, as it is a metabolising enzyme and belongs to the CYP group (molecular weight 56,849 Da). Amounts of 100, 50, 10 and 1 pmol were run on a gel (Figure 3.3). Except for the 1 pmol-sample a band for all the CYP2E1 recombinant preparations was visible on this gel using a Coommassie stain, providing an estimate for the amount of protein. All bands were cut out and digested including a band in the 1 pmol-lane (not shown), which used the other bands as markers to cut out the appropriate region. A tenth of each re-dissolved peptide extract was injected, thus between 10 and 0.1 pmol was submitted to the LC-MS system. The MS analysis of these samples could detect the recombinant protein for each gel loading, including the 1 pmol sample (100 fmol injected), which was assigned a Mascot score of 95 and sequence coverage of 17.5%. According to the Mascot search 3 peptides could be detected. When 10 pmol of recombinant protein was loaded onto the gel (1 pmol injected), the sequence coverage increased to 28% with 15 peptides identified. For the 50 pmol sample the sequence coverage was 44% (22 peptides) and 49% for the 100 pmol sample (24 peptides). Full data information about protein identification can be found on the accompanying CD (file 2E1_experiment). In this experiment there was only one protein loaded and also the digested peptide mixture for MS contained the recombinant CYP2E1 peptides. The only other detectable protein in the sample was keratin as a main contamination present in all samples. Lower amounts of CYP2E1

were not tested, as this was difficult due to the detection limits of the Coommassie stain, and since as this experiment was designed to prove the ability of the method to detect CYPs, which was successful for recombinant CYP2E1.

Table 3.3: Recombinant CYP2E1 in-gel-digestion control results.

Protein on gel [pmol]	Protein score	Protein sequence coverage [%]	Number of identified peptides
1	95	17	3
10	418	28.4	15
50	1927	44.2	22
100	2732	49.3	24

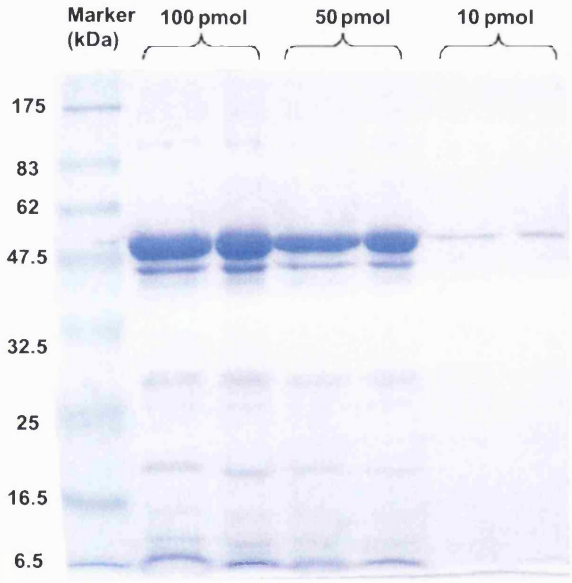


Figure 3.3: SDS-PAGE of recombinant CYP2E1. 100, 50 and 10 pmol of recombinant CYP2E1 were applied onto the gel (reduced with DTT on the left-hand side, reduced with DTT and alkylated with iodoacetamideand on the right-hand side) analysed following the protocol described in Experimental. The band for CYP2E1 can be seen at approximately 47.5 kDa. Other proteins visible are presumably contaminants or degradation products.

3.5.2 Validation with liver sample

The experiment with recombinant CYP2E1 showed that using the described LC-MS/MS method it is possible to detect CYP2E1 to a level of 1 pmol on gel or 100 fmol on-column. CYPs were detected with this method previously in liver microsomes using the LCQ mass spectrometer, so the next step was to ascertain that CYPs in human liver

could be detected with the QTOF. A liver sample, which had been analysed previously using the LCQ and known to contain CYPs, was used as a control sample [245]. Thus, the microsomal preparations of this human liver and one lung microsomal sample (P4) for comparison were run in parallel on a gel (Figure 3.4) and then analysed using the same protocol. Table 3.4 lists the identified CYPs in the human liver sample. Further information about all proteins identified can be found on accompanying CD (file liver_experiment). In the liver sample a total of ten CYP enzymes were detected and these enzymes had also been identified previously in this liver sample (data not shown). However, using this method, only one CYP, namely CYP5A1 or thromboxane-A synthase 1, was identified in the lung sample of individual P4 as described above (3.4). These results show that the method is capable of detecting CYP proteins in liver microsomes, however further changes of the method might be necessary to detect these proteins from lung to improve sensitivity and allow the detection of lower abundant CYPs. As there are CYP proteins present in human lung [162], the question remains whether the method developed here can be improved to be more sensitive and detect lower abundance proteins than the ones identified so far.

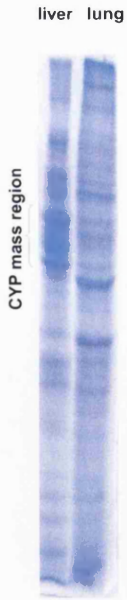


Figure 3.4: Comparison of liver and lung microsomes on SDS-Gel. Microsomal preparations of human lung and liver were processed following the same protocol. The resulting gel shows darker bands in the 45 to 62 kDa area in the liver sample than in the lung one. Lung microsomes derived from individual P4.

Table 3.4: Identified CYP enzymes in human liver microsomes. A liver microsomal preparation was treated in a similar manner to lung samples and the same protocol was applied. Individual peptide score had to be over 35 for acceptance of data; except for CYP4F3 all proteins were assigned by at least 2 peptides. Sequence coverage is given as percentage of whole sequence assigned with experimental hits.

CYP enzyme	Mascot protein score	Protein coverage [%]	Number of peptides identified (unique peptides)
1A2	189	14.9	8 (5)
2A6	199	13	5 (1)
2A13	150	7.3	3 (0)
2B6	57	3.5	2 (1)
2C8	101	6.9	3 (1)
2C9	106	7.3	3 (1)
2E1	72	8.9	2 (2)
3A4	60	5.4	2 (1)
4A11	59	5	2 (0)
4F3	38	2.7	1 (0)

3.6 Method improvement

3.6.1 Introduction of smaller *m/z* ranges

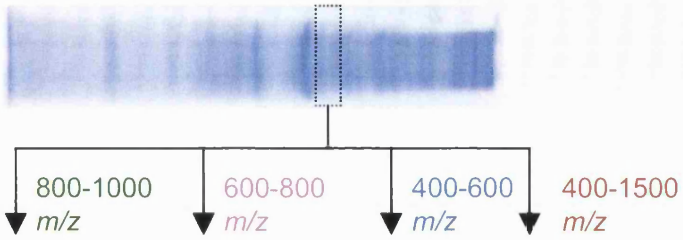
At this point it was decided to try to improve the sensitivity of the method. The data-dependent acquisition algorithm selected only peptides of greatest abundance to perform a MS/MS analysis. To obtain information from less abundant peptides (and proteins), the MS survey scan prior to MS/MS scans – which is used by the data-dependent acquisition algorithm to select ions of interest – was divided into three mass ranges, thus scanning from *m/z* 400 to 600, 600 to 800 and 800 to 1000. That means one sample was injected three times, each with a different *m/z* range for the MS survey scan. These mass ranges were chosen according to the experiment described in 3.4, where of 3626 experimentally detected peptide masses only 30 showed a *m/z* higher than 1000. This way the *m/z* range was reduced, but the run time was kept the same,allowing more parent ions from one sample to be selected by DDA. The submission of samples to three small *m/z* ranges is now referred to as method 2.

3.6.2 Application of smaller *m/z* ranges to CYP region

The impact of using three smaller *m/z* ranges instead of one large *m/z* range in the MS survey scans is shown in Figure 3.5. One gel band of sample P3 was submitted to both methods and thus injected four times. The initial method, scanning over a larger *m/z* range (400–1500 *m/z*), identified 5 proteins. As visible from the table in Figure 3.4, a higher number of proteins were identified when scanning the smaller *m/z* ranges separately. The sum of all proteins detected with method 2 was 51, which makes it a ten times higher output of identifications for this band. More details about identified proteins can be found on the enclosed CD.

3.6.3 Application of smaller *m/z*-ranges to a whole gel-lane

As described in 3.6.2, using method 2 there was a much higher number of proteins found than just using the method 1. For the sample described, there was a 10 fold difference in the methods. This was an interesting finding and the following experiment should show whether this improvement could be reached for all other bands. Therefore the total lane of individual sample P3 was analysed with method 1 and method 2. This was a way to look at the whole microsomal lung proteome and compare the two methods in the whole mass range of the gel. For that reason the gel lane was cut into 36 bands (Figure 3.6) and each was submitted to in-gel digestion and LC-MS/MS analysis using methods 1 and 2. As one band was injected four times, this resulted in 144 LC-MS/MS runs altogether. For data processing of Mascot search results, the ProteinCenter software was used. Thus, a total identification of 304 proteins with method one compared to 473 proteins with method two was achieved (Table 3.5), showing a 1.5 times increase in the number of proteins identified with method 2. Among the three different *m/z*-ranges, most proteins could be found using the medium mass range (600-800 *m/z*), followed by the 400-600 *m/z* range. This indicated that tryptic peptides were preferably found in these *m/z* regions (400-800 m/z). For detailed information of proteins in the different *m/z* ranges, see accompanying CD.



Actin	Epoxide hydrolase 1	Epoxide hydrolase 1	78 kDa glucose-regulated protein precursor
Dolichyl-diphosphooligosaccharide-protein glycosyltransferase	Synaptic vesicle membrane protein VAT-1 homolog	Actin	Calreticulin precursor
Serpin H1 precursor	Sulfide:quinone oxidoreductase	Adipocyte plasma membrane-associated protein	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase 67 kDa subunit
Alpha-enolase	Serpin H1 precursor	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex
Synaptic vesicle membrane protein VAT-1 homolog	ATP synthase subunit beta	Sulfide:quinone oxidoreductase	Advanced glycosylation end product-specific receptor precursor
Beta-enolase	Protein disulfide-isomerase A6 precursor	Serpin H1 precursor	
Flotillin-1	Cytochrome b-c1 complex subunit 1	Tubulin alpha-1A chain	
Annexin A7	Adipocyte plasma membrane-associated protein	Serum albumin precursor	
Tubulin alpha-1A chain	Arylacetamide deacetylase-like 1	Elongation factor 1-alpha 1	
Eukaryotic translation initiation factor 3 subunit F	Actin	Eukaryotic translation initiation factor 3 subunit F	
Epoxide hydrolase 1	Trifunctional enzyme subunit beta	Trifunctional enzyme subunit beta	
Arylacetamide deacetylase-like 1	Alpha-enolase	Protein disulfide-isomerase A3 precursor	
Thioredoxin domain-containing protein 4	EH domain-containing protein 2	Tripeptidyl-peptidase 1 precursor	
Sulfide:quinone oxidoreductase	Endoplasmin precursor	Vimentin	
Serum albumin precursor	Flotillin-1	Sister chromatid cohesion protein PDS5 homolog A	
EH domain-containing protein 2	Eukaryotic initiation factor 4A-I	EH domain-containing protein 2	
Nuclease sensitive element-binding protein 1	Flotillin-2	Alpha-enolase	
Tubulin beta chain	Cathepsin D precursor	Protein disulfide-isomerase precursor	
Protein disulfide-isomerase A6 precursor	Protein disulfide-isomerase A3 precursor	Flotillin-2	
Calnexin precusor	Ig gamma-1 chain C region	Membrane copper amine oxidase	
Adipocyte plasma membrane-associated protein	Pulmonary surfactant-associated protein D precursor	Cytochrome b-c1 complex	
78 kDa glucose-regulated protein	CD63 antigen	Arylacetamide deacetylase-like 1	
	Tripeptidyl-peptidase 1 precursor	Calnexin precursor	
	Serum albumin precursor	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase	
	Ribonuclease inhibitor	Ezrin	
	Proliferation-associated protein 2G4	Endoplasmin precursor	
	Advanced glycosylation end product-specific receptor	Tapasin precursor	
		Development and differentiation-enhancing factor 2	
		Synaptic vesicle membrane protein VAT-1 homolog	
		XK-related protein 3	
		Iduronate 2-sulfatase	
		Eukaryotic initiation factor 4A-I	

Figure 3.5: Comparison of protein identification using a single run or three runs with different *m/z* ranges. The microsomal fraction of individual P3 was reduced, alkylated, digested and separated using SDS-PAGE. Band 16 was cut, extracted, digested and analysed using LC-MS/MS. The proteins were identified using Mascot

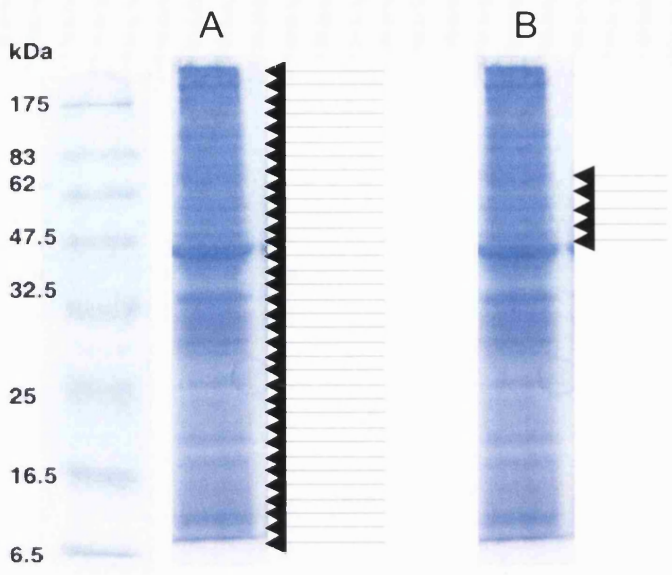


Figure 3.6: Two different approaches were used to cut out samples. Either the whole gel was cut in 36 parts of approximately equal size (A) or the five bands in the mass range of 45 to 62 kDa were cut (B).

Table 3.5: Overview of the number of protein hits for each method. The percentage of single peptide hits proportion is given. Total number of proteins identified with method 2 is stated as well as identification in each single mass range-run.

	Method 1	**Method 2**	**Method 2**		
			400-600 *m/z*	**600-800 *m/z***	**800-1000 *m/z***
Number of proteins	304	473	297	347	169
Single peptide hits [%]	37	19	31	38	27

3.7 Comparison and organisation of data

The results so far have shown (3.6), that the splitting of the MS-survey scans into smaller *m/z* ranges had the advantage of identifying more proteins. In the case of a whole lane, i.e. data of 36 bands, a 1.5 fold increase in protein identifications could be achieved. The comparison of the individual data sets via ProteinCenter software could give more detailed information. The numbers stated in Table 3.5 only give the total numbers, but it is necessary to discover what the differences in both methods were.

The comparison of the data sets showed that a total of 543 proteins were listed from both searches and 208 of these are common to both. That means that there were 96 proteins detected by method one which were not detected with method two, but 265 more proteins that were detected with method two and not with method one. That implied that method two was better in the sense of output of identifications, yet some proteins were not detected.

3.8 Organisation of data into subgroups

Data could be arranged according to different gene ontology (GO) criteria; either in which biological processes the identified proteins were involved, which molecular functions they had or from which cellular component they originated. The following statistics were produced for the data derived from method two for individual P3, so a total of 473 proteins were sorted. As the samples derived from lung microsomes, it was interesting to see how many proteins were assigned to this compartment after analysis.

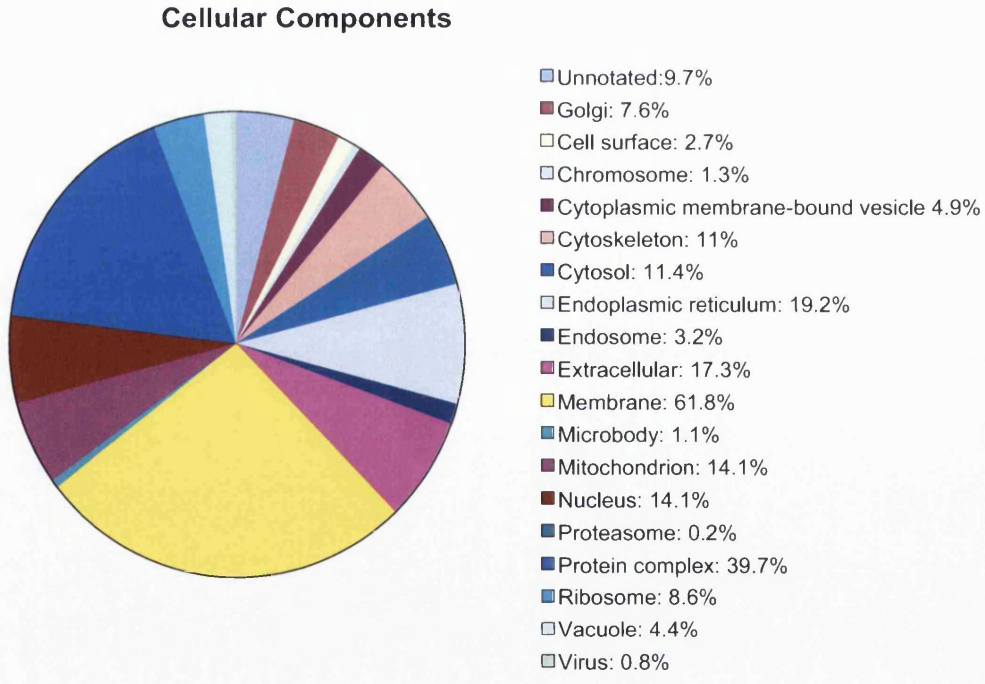


Figure 3.7: Overview of proportion of proteins in cellular components for individual P3, analysed with method 2. (One protein can be associated with more than one component, so the sum of all percentages can be higher than 100%.)

As Figure 3.7 shows, the biggest fraction was assigned as membranes with 61.8%. The endoplasmic reticulum, which is part of the microsomal fraction showed the third highest percentage (19.2%). Although microsomes were not listed as a fraction on its own, estimations for this component could be made by looking at these two compartments. Also membranes of the Golgi apparatus could be contained in microsomes, so all three compartments together show a high proportion (61.8%+19.2%+7.6%). It was also noticed that mitochondrion got a relatively high percentage (14.1%) as well as nucleus (14.1%), both fractions which should have been separated through subcellular fractionation.

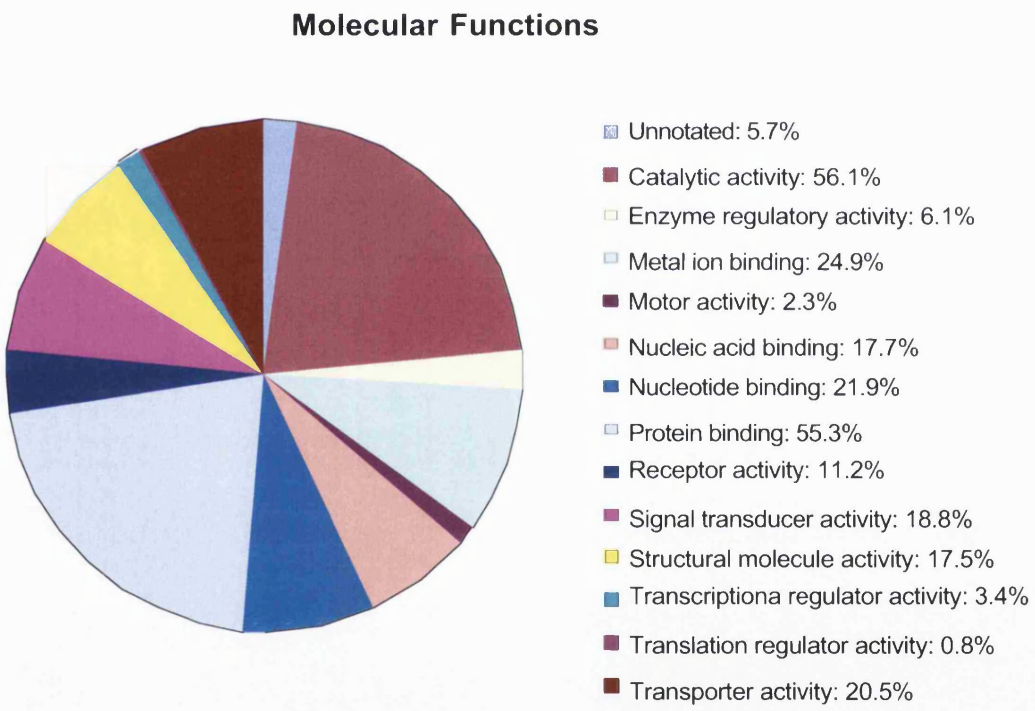


Figure 3.8: Overview of proportion of proteins by molecular function for individual P3, analysed with method 2. (One protein can be associated with more than one component, so the sum of all percentages can be higher than 100%.)

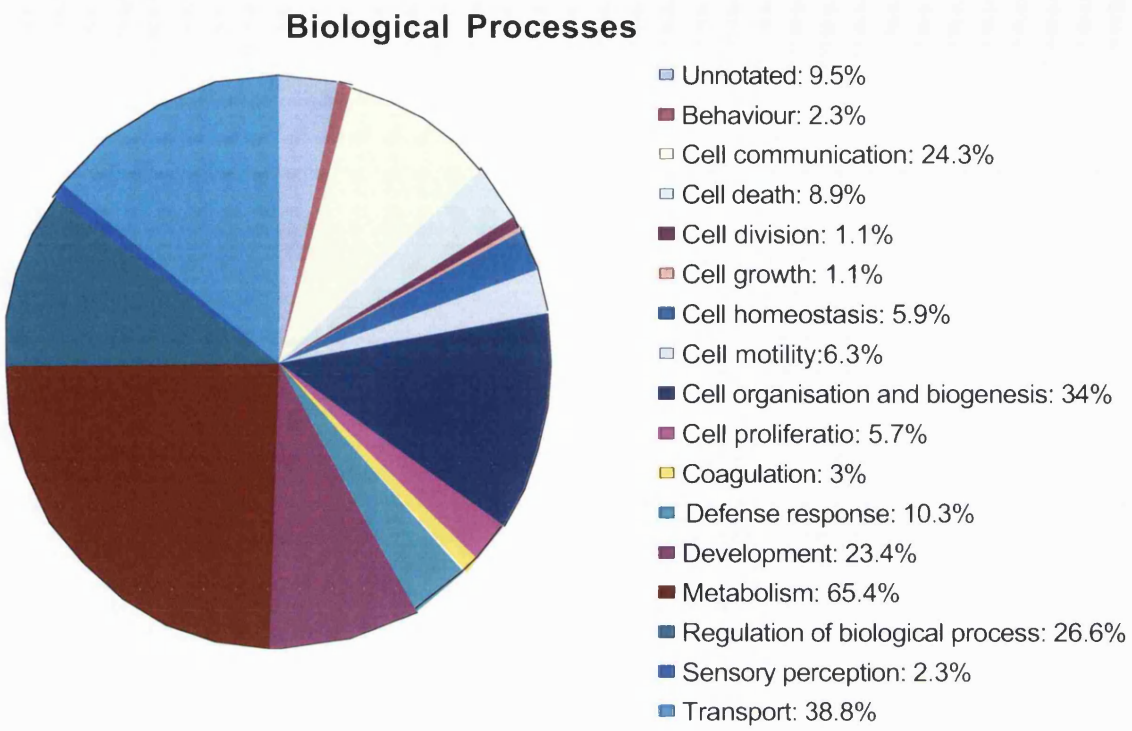


Figure 3.9: Overview of proportion of biological processes assigned to proteins identified for individual P3, analysed with method 2. (One protein can be associated with more than one component, so the sum of all percentages can be higher than 100%.)

The GO molecular functions (Figure 3.8) showed highest percentage of 56.1% of proteins exhibiting a catalytic activity, which is expected in the case of microsomal samples as it is supposed to contain many enzymes able to catalyse biochemical reactions. Any protein showing interactivity with other proteins or protein complexes presented the second highest part (55.3%). Also proteins with transporter activity could be expected in microsomes, as in the endoplasmic reticulum proteins are synthesised and exported into other compartments; this group of proteins was here assigned 20.5%.

When all identified proteins were sorted according to their biological processes, the biggest proportion was assigned with 65.4% to metabolism followed by transport with 38.8% and cell organisation and biogenesis (34%). This could all be applied to processes in the microsomes, as metabolising enzymes are present in the endoplasmic reticulum. Synthesis of proteins and their transported into other parts of the cell also takes place here.

Similar ratios were also obtained for protein identifications deriving from the application of method 1 for individual P3 (data not shown).

3.9 Application of method 2 to sample P4

Subject P4 was also analysed with method 2, so the whole lane was used, submitting 36 samples to three LC-MS/MS injections each; method 1 was not applied to the whole lane. In total 288 proteins were identified, of which 191 were found in sample P3 as well. 97 proteins were only found in sample P4. For sample P3 in total 473 proteins were found using method 2, so 185 proteins more compared to sample P4. A detailed overview of protein identification is given in appendix A (table ii). As only two individuals were analysed using method 2, it is difficult to draw any conclusions about differences in these samples. It cannot be stated by this what number of proteins or which proteins would have to be expected. In order to do this, more individuals would have to be tested using method 2 and then data would have to be compared.

Variation was also observed for samples P1, P2, P5, P6, P7 and P8, when the CYP area only was analysed using method 1 (3.4). The individual analysis gave following differences in the number of identified proteins. For P1 a total of 79 proteins, for P6 67, for P7 75 and for P8 67 proteins were found. For samples P5 46 and for P2 only 36 proteins were found. However analysis of the six individuals was only performed for five bands of the whole gel, so it might be possible that there were some inter individual variations, but for significant conclusions whole gel-bands would have to be analysed.

3.10 Resume and discussion

The method applied in this work, consisting of 1D gel separation, protein digestion and following LC-MS/MS detection of resulting peptides has resulted in a wide range of identified proteins. In sample P3 a total number of 473 individual proteins were identified. A list of all the proteins can be found on the enclosed CD. With a division of the normal MS/MS run (method 1) into smaller m/z ranges (method 2) it was possible to increase the number of identified proteins. Using method 1 only would have resulted in 304 proteins, whereas with method 2, 169 more proteins were identified. There were 96 proteins, which were uniquely identified with method 1. Comparing the three m/z ranges which were applied in method two it was observed that most proteins were detected in the first two m/z ranges, and the last one (m/z 800–1000) contributed less to

the identification. Nevertheless the subdivision into smaller *m/z* ranges was a successful means to increase the number of identifications. Method 2 yielded a 1.5 fold higher output of proteins (473 in method 2 compared to 304 with method 1). The identification of proteins out of the complex tissue sample also benefited from the multiple separations applied to the sample. The first step was carried out by ultracentrifugation to obtain the microsomal fraction. The proteins in the samples were then subjected to size-dependent separation in 1D-SDS-PAGE. Peptides were separated in the LC and then gas-phase fragmentation in the MS scanning for three different *m/z* ranges was the ultimate separation step.

The results so far show that the described LC-MS/MS method is also suitable for the detection of metabolising enzymes. Using recombinant CYP enzymes as standard for important metabolising enzymes, it could be shown that an amount of 1 pmol on-gel (100 fmol on-column) is detectable.

However using this method it was not possible to detect CYPs in lung tissue, other than CYP5A1. There is more than one explanation for this. A possible explanation could be that the concentration of the lung CYPs may be below the limit of detection or other proteins in the sample may have been too abundant. The presence of high levels of keratins in the sample can also have a significant effect on the quality of the data. During processing, samples can become contaminated by keratins from the environment. In all working steps samples are exposed to possible sources of contaminations: skin, hair, nails, clothes etc., as well as keratins present in the air or surfaces and dust. The contamination can cause a problem if their abundance is higher than any other proteins. As in the data dependent mode the instrument is selecting the most abundant peaks for fragmentation, peaks representing keratin-peptides were then primarily detected. The peptide peaks of lower intensity belonging to low abundance proteins were withdrawn from detection. However, a strict cleaning regime and wearing protective clothing did not prevent keratin contamination. This shows the importance of keratin-free environment, which is crucial when proteins of interest are expected to be in low abundance.

The difficulties in detecting CYPs could also be due to the properties of the lung itself. The liver as a place of high metabolism was used in this and previous works as a useful sample for detecting metabolising enzymes. Although the pulmonary metabolic

capacity is far lower than that in liver [246], it must be taken into consideration, as the large vascular surface of the pulmonary veins is exposed to circulating xenobiotics in the blood. Results achieved so far, show that a large number of different types of proteins are detectable with this method (3.6.3). Comparing lung and liver tissue results, only one CYP was identified in the lung samples (3.5.2). Additional experiments however proved the method to be suitable to detect recombinant CYP standards (3.6.2) as well as CYP in liver tissue (3.5.2). As these enzymes are of low abundance in the lung, in fact only 10% of the liver content, as quoted by Zhang *et al.* [162], this leads to the assumption that the applied methods 1 and 2 are not sensitive enough to detect them.

Also the nature of the sample has to be considered. The lung as an organ is consisting of 40 different cell types [4] with differing amounts of CYPs, which makes the identification in a complex sample such as microsomes difficult. Using spectroscopic quantification, it was shown earlier [235, 247] that the amount of microsomal CYP protein is about 2–10 pmol/mg protein. In earlier work by McManus *et al.* CYPs could not be detected in human lung [246].

In experiments reported here, the amount of total microsomal proteins loaded on a gel was 25 µg, which would correspond to 50–250 fmol CYP proteins. Assuming a mean molecular weight of 65,000 Da for a protein, the number of moles would correspond to 385 pmol protein added per lane on the gel, including all types of proteins present in the sample and not only metabolising enzymes. This implies a several thousand-fold higher ratio of total proteins compared to CYPs. This is a clear indication of the low abundance of this protein class in lung tissue and a possible cause for non-detection.

Better detection of various CYP proteins in human lung microsomes on the protein level was achieved so far when performing immunochemical studies, such as a western blot [234, 235, 248], which requires specific monoclonal antibodies. Knowing that CYPs are localised in different amounts in the different lung-cell types, it could be a future goal to separate out e.g. clara cells [249, 250], high in CYP, and apply the LC-MS/MS method on a microsomal preparation of these samples including immunochemically comparing both methods. In order to detect specifically CYPs, also the method of multiple reaction monitoring [156] would be suitable to scan selectively for particular masses and control their fragmentation into expected product ions. Thus LC-MS/MS could be

applied and no monoclonal antibodies need to be used. This will be attempted in the following chapters, where methods for the quantifications of CYP1A2 and CYP2E1 are developed.

Chapter 4 Quantification of Cytochrome P450 2E1

4.1 Introduction

Cytochrome P450 2E1 (CYP2E1) contributes approximately 7%–9% to the pool of cytochrome P450 enzymes of human liver [216, 217] and plays an important role in a wide range of metabolic processes (see 1.5.2.4). It takes part in the metabolism of a large variety of substrates including aromatic compounds (e.g. acetaminophen, toluene, benzene, chlorzoxazone), alcohol (ethanol), halogenated anaesthetics (e.g. halothane), heterocyclic compounds (e.g. isoniazid, pyridine), some nitrosamines (e.g. *N*,*N*-dimethylnitrosamine) [208] and also endogenous compounds such as acetone and fatty acids (linoleic and arachidonic acids) [218]. In turn, expression of CYP2E1 can be induced by many of these substances. An interest in the function of CYP2E1 leads to questions of its occurrence and quantities in various tissues and body fluids. There is a need to assess its abundance so as to draw conclusions about its importance in biological processes such as metabolism and disease development. Quantification of CYP2E1 has been performed using RT-PCR [211], yet these results only provide information about transcription levels and mRNA content but not the cellular protein content. Cellular protein content depends not only on mRNA translation but also on protein degradation and these processes can be regulated independently. For this reason, the direct quantification of the protein provides more accurate information on protein abundance and is preferable in this regard. Quantitative analysis on the protein level can be performed by Western blot [212], but this is dependent on the availability of specific antibodies, and may not assure 100% selectivity; furthermore, only specific CYP isoforms can be quantified per analysis. Liquid chromatography (LC) with tandem mass spectrometry (MS/MS) is commonly used for the quantification of small molecules, like drugs and metabolites, and is also gaining popularity for qualitative and quantitative

analysis of proteins [214]. For CYP analysis LC-MS/MS has been used primarily in a discovery mode, where the goal is protein identification [243, 251]. Although LC-MS/MS has been used for protein quantification over the last decade, this application still remains challenging. Currently there are a variety of LC-MS/MS methods being used in proteome science for both absolute and relative protein quantification [252, 253]. Among these methods, the use of stable isotope labelled tryptic peptides offers a promising method of absolute quantification [155] [156]; briefly, proteins are digested with trypsin at which point a stable isotope labelled tryptic peptide is added to the mixture. The isotope labelled peptide is chosen to have an amino acid sequence unique to the protein of interest but to contain a heavy isotope labelled amino acid. Quantification is then performed by comparing the mass spectrometric ion-current of the stable isotope labelled peptide with its native analogue generated by digestion of the target protein. Assuming tryptic digestion of the target protein is complete, the abundance of the target protein can be inferred from that determined for its tryptic peptide. The use of stable isotope labelled peptides has the advantage of allowing absolute protein quantification with no additional experimental steps other than addition of the labelled compound and can easily be adapted to an established in-gel digestion method [243, 251].

Here, a method for the quantification of CYP2E1 in human liver tissue, which contains a large number of CYPs in high abundance, has been developed, including method validation using authentic standards.

4.2 Peptide properties and fragmentation pattern

For the quantification of the protein, a peptide of amino acid sequence unique to CYP2E1, i.e. GTVVVPTL*DSVLYDNQEFPDPEK, was used. The peptide was labelled with six ^{13}C and one ^{15}N at leucine-8 (L*), resulting in a mass difference between native and labelled peptide of 7 Dalton. The monoisotopic molecular masses of the peptides were 2561.25 Da and 2568.25 Da, respectively. Solubility properties of the peptide and recombinant protein were tested. As the protein was not soluble in 0.1% formic acids, standards were prepared in 0.1% formic acid, containing 100 mM ammonium bicarbonate and 5% acetonitrile and stored in 11.7 pmol/µl aliquots (-20°C).

To determine suitable precursor and product ion pairs for MRM analysis, mass spectra and product ion spectra (MS/MS) of the labelled peptide were recorded. The peptide was introduced by direct infusion into the nano-spray source of the LCQ and the instrument conditions optimised to achieve maximum sensitivity. Figure 4.1 shows a typical mass spectrum of a 100 fmol/µl solution with the double charged peak ($[M+2H]^{2+}$) at *m/z* 1285.3 and the triple charged peak ($[M+3H]^{3+}$) of lower intensity at *m/z* 857.4. Additionally, the double charged peptide provided more intense fragment-ion signals than the triple charged ion (area and signal to noise for double charged fragment *m/z* 1285.3 → 1057.6: 3.1×10^6 au, S/N 17; area and signal to noise for double charged fragment *m/z* 857.4 → 1057.6: 1.3×10^6 au, S/N 3). Therefore the double charged ion was selected as the parent ion to monitor fragmentation in MS/MS. The most abundant fragment ions in the MS/MS spectra of the $[M+2H]^{2+}$ ion of the native and labelled peptide were at *m/z* 1054.3 and *m/z* 1057.6, respectively. Both represent the y_{18}^{2+}, resulting from cleavage of the peptide *N*-terminal to proline-6 (fragmentation pattern shown in Figure 4.2). These fragment ions were optimal for MRM analysis as they contain leucine-8 allowing the separation of labelled and non-labelled peptide fragments by mass difference. No additional fragments between y_{16} and y_{22} were detected, thus the y_{18}^{2+}-fragment was chosen for MRM. Monitored masses are shown in Table 4.1. For quantification, the ion transitions *m/z* 1281.8 → 1054.3 (analyte) and *m/z* 1285.3 → 1057.6 (standard) were used. The isolation window was 1.5 *m/z* units either side of the mass chosen.

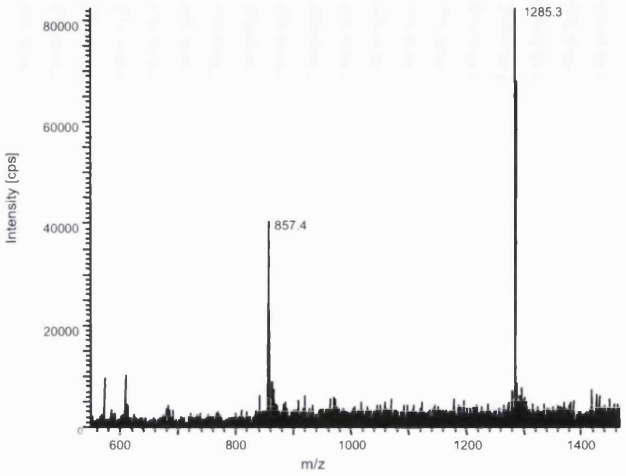


Figure 4.1: Mass spectrum of the isotope labelled internal standard peptide for the quantification of CYP2E1 (GTVVVPTL*DSVLYDNQEFPDPEK; direct infusion, 100 fmol/µl). Double (1285.3 *m/z*) and triple charged (857.4 *m/z*) peaks are shown.

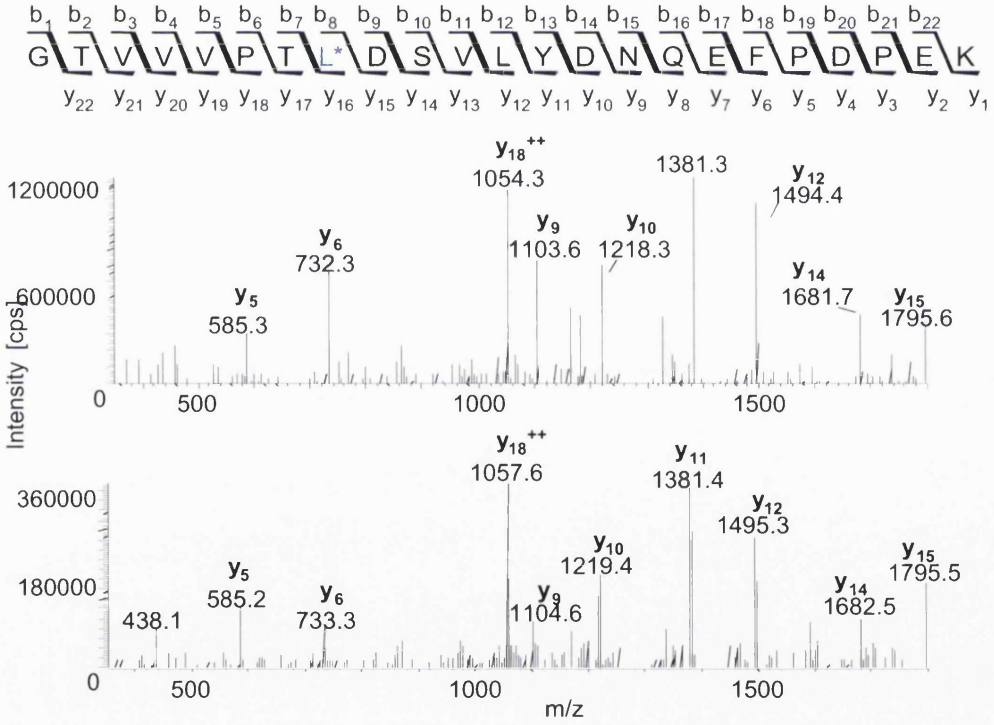


Figure 4.2: Mass spectra and amino acid sequence of the native (upper spectrum) and labelled internal standard peptide (lower spectrum) with the indicated possible y- and b- ions. In both spectra the precursor ion was the $[M+2H]^{2+}$ ion. For MRM experiments, the fragmentation of the $[M+2H]^{2+}$ into the y_{18}^{++} fragment was monitored (1281.8 → 1054.3 and 1285.3 → 1057.6).

Table 4.1: Mass, ions and fragments of native and labelled internal standard peptide (molecular mass is given in g/mol, ions in *m/z;* theoretical values are monoisotopic).

Peptide	Molecular mass	theoretical $[M+2H]^{2+}$	observed $[M+2H]^{2+}$	observed *m/z* of y_{18}^{2+} fragment
Native	2561.3	1281.6	1281.8	1054.3
Labelled	2568.3	1285.1	1285.3	1057.6

4.3 LC Method development

To separate and concentrate the sample, a capillary LC system (see experimental) was used. A mixture of recombinant CYP2E1 digest and labelled peptide eluted from the column after approximately 40 minutes (Figure 4.3). This was very important, as the quantification method relied on the co-elution of both peptides. The hydrophobic character of the peptide explains the late elution of the peptide, as at about 37 minutes the ratio of organic solvent was at its highest and was used to wash off any remaining hydrophobic substances from the column.

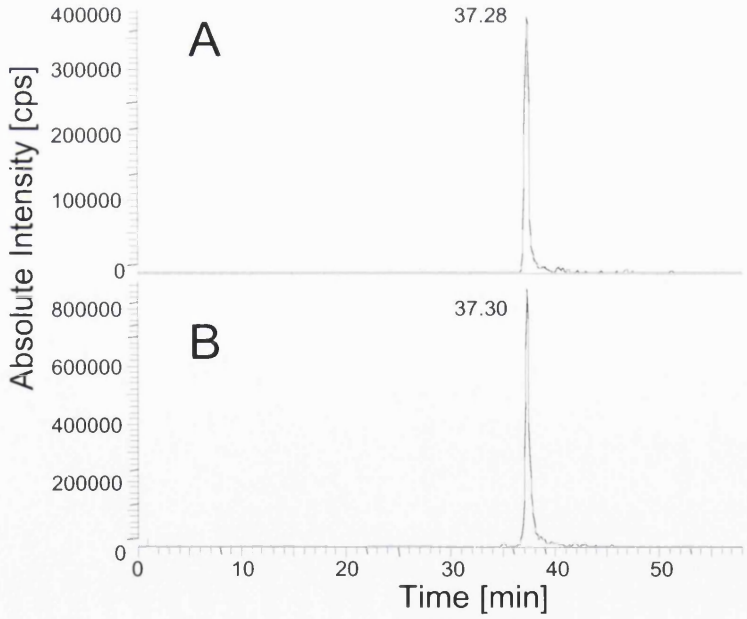


Figure 4.3: The reconstructed ion chromatogram of native and labelled internal standard peptide eluting at about 37.3 minutes. For the native peptide the transition 1281.8 → 1054.3 (A) and for the labelled internal standard peptide 1285.3 → 1057.6 (B) were monitored. The mobile phase composition and gradient are described in section 2.3.2.2

4.4 Validation of the quantification method

After determination of the suitable transitions for the MRM experiment, the method had to be validated before performing reliable quantification in tissue samples. Aspects such as reproducibility, sensitivity, accuracy and precision had to be tested. The accuracy results will describe how close the experimental value comes to the expected one and precision results will show variations in individual data. The most suitable way to do this is using standards as similar to the analyte as possible. Instead of the endogenous form of the protein, recombinant CYP2E1 was used. For the analysis of actual liver microsomes, these were separated by 1D-SDS PAGE, followed by an in-gel digestion, extraction, and then the CYP2E1 specific MRM method was used for LC-MS/MS analysis; the same method was applied to the recombinant protein. As in-gel digestion may result in loss of protein due to incomplete extraction from the gel, an in-solution digest was also performed. The matrix – liver microsomal preparation – was also tested for interference by spiking liver microsomes with recombinant CYP2E1 into liver microsomes. To ensure a robust and reproducible method, interferences with the gel-matrix and carry-over in the LC-MS runs were also investigated.

4.4.1 In-solution digest calibration

The first experiment was conducted with an in-solution digest of recombinant CYP 2E1 mixed with labelled-peptide. In this way there should be no extraction-losses, as would be encountered during an in-gel digestion, and close to 100% recovery of protein should be expected. The in-solution digest was prepared as described in the methods section (2.2.9), aliquoted and stored at -20°. The stock had a concentration of 0.2 µg/µl CYP2E1, corresponding to 3.67 pmol/µl. To this stock solution, different amounts of labelled peptide were added to achieve a ratio (mol/mol) of 1 to 1, 1 to 5 and 5 to 1. The mixtures were then diluted, injected and analysed using the 58 min LC-MS-MRM method (2.3.2.2), monitoring for the described fragment ions (*m/z* 1281.8 → 1054.3 and 1285.3 → 1057.6). The ratios (unlabelled/labelled) of the areas under the curves in the reconstructed ion chromatograms for the respective transitions were determined and compared with the theoretical ratio. The theoretical or expected ratio can be calculated, as the amounts of recombinant protein digested as well as the amount of labelled

peptide added are known. A plot of the theoretical ratio versus the experimental ratio resulted in the graph shown in Figure 4.4.

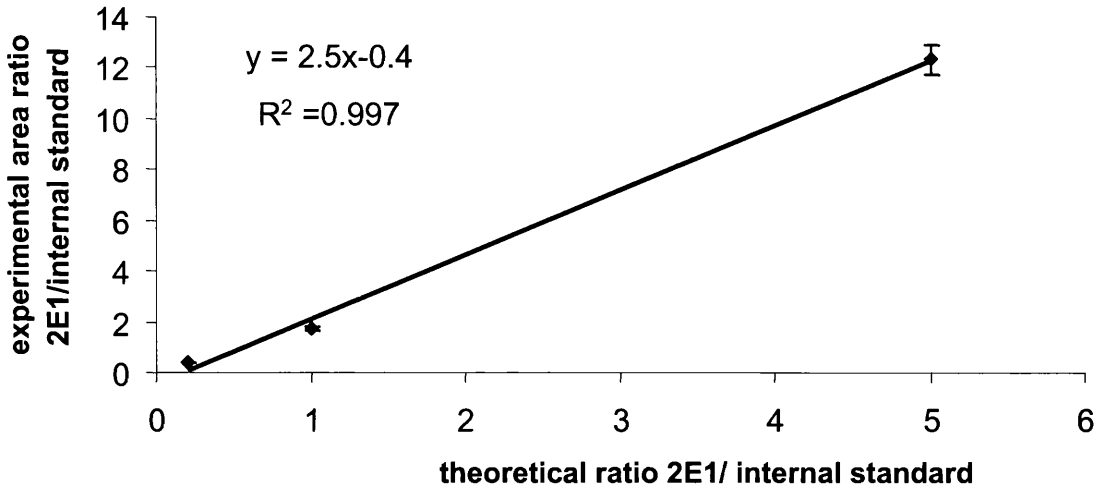


Figure 4.4: In-solution calibration experiment of CYP2E1. Theoretical ratios (ratio of amounts of native and labelled peptide) are plotted versus experimentally measured area ratios. Error bars are standard error of the means.

The R^2 value (R=correlation coefficient) is used as an indicator for a linear relationship between theoretical and practical data. Here, $R^2 = 1$ indicates linearity, so 0.9971 is an acceptable value for R^2. The variability of the method, shown by CV-values of 12% to 14 % (Table 4.2) were also well within the range recommended for bioanalytical methods by the FDA (< 15% except for limit of quantification) [254].

Table 4.2: In-solution calibration experiment. Average experimental ratio (analyte/internal standard peptide areas) is used to calculate the injected 2E1 amount. This is compared to the injected amount and recoveries are given in %. N is the number of accepted MRM runs of a particular concentration.

CYP2E1 digest injected [pmol]	Labelled peptide injected [pmol]	Theoretical ratio native/labelled	N	Average experimental area ratio native/labllled	Stdev	SEM	CV [%]	2E1 calculated (pmol)	2E1 recovery [%]
0.38	0.38	1	7	1.72	0.24	0.09	14	0.58	172
0.2	1	0.2	8	0.38	0.05	0.02	12	0.38	190
1	0.2	5	8	12.31	1.73	0.61	14	2.46	246

The recovery rates found ranged between 172% and 246% (also visible from the slope of 2.5), which is very high and indicates a 2-fold excess of native to labelled peptide, suggesting an incomplete recovery of the labelled peptide due to low solubility in the digestion buffer. Recovery rates and the effect of digestion will be discussed in more detail in the next chapter (chapter 5).

The linear equation determined in the calibration experiment for the in-solution digest (y=2.5x-0.4, Figure 4.4) can be used to recalculate the actual ratio, by using the experimentally determined ratio as y and calculating the theoretical ratio as x. Thus, the new calculated ratios are also used to recalculate the theoretical amount of CYP2E1 injected as well as the recovery rates; the recovery for 1 to 1 and 5 to 1 ratio's are 85% and 101% respectively (Table 4.3). A possible explanation for a slope larger than one – as would be expected – is the incomplete solubilisation of the peptide due to its hydrophobicity.

Table 4.3: Recalculation of protein content and recovery of calibration samples (Table 4.2) using the calibration line for in-solution digest (Figure 4.4, †y=2.5x – 0.4).

CYP2E1 digest injected [pmol]	Labelled peptide injected	Theoretical ratio native/labelled [pmol]	N	Calculated area ratio (with equation)†	stdev	sem	CV [%]	Recalculated 2E1 [pmol]	2 E1 recovery [%]
0.38	0.38	1	7	0.85	0.24	0.09	28	0.32	85
0.2	1	0.2	8	0.32	0.05	0.02	15	0.32	162
1	0.2	5	8	5.03	1.73	0.61	34	1.01	101

4.4.2 In-gel-digest calibration

As samples were to be quantified from in-gel digestion, the effect of the in-gel digestion and gel extraction was investigated. Recombinant CYP2E1 was treated and run in the same way as liver samples, using 1D-SDS-PAGE including reduction and alkylation of the protein, and in-gel-digest with trypsin in the presence of labelled peptide. Samples were analysed as described above. Different amounts and different ratios of the peptides were used again in order to show a linear relationship between expected and observed values. After in-gel digestion, peptide extracts were dried and re-suspended in 50 µl solvent and 5 µl of it were injected. Table 4.4 shows the injected amount, the experimental ratios of light to heavy peptide and the calculated injected 2E1 amount using the experimental ratio and the known amount of labelled peptide added. The standard deviations as well as the CV-values were acceptable; only one CV value was above 20% (29%). However, the recoveries for the samples with a low amount of recombinant CYP2E1 loaded on gel (5 pmol) were very low (23% and 36% only) whereas the samples with higher recombinant protein showed recoveries between 98 and 122%. Figure 4.5 shows the plot of theoretical versus measured ratios. The R^2 of 0.973 is not as good as for the in-solution calibration, however, the plot suggests a non-linear relationship, which could be due to instrumental limitations, e.g. ion-trap over-fill or detector response. A better fitting curve for connection of points was indeed found when applying a quadratic equation for regression analysis. In fact the correlation coefficient was 0.999 (Figure 4.6). Also the recoveries improved (Table 4.5), so for this

particular experiment a quadratic equation for regression analysis would be suitable. The influence on the calculation of CYP2E1 content in biological samples using the quadratic equation was investigated as well (4.6.5).

Table 4.4: Overview of in-gel digestion calibration. Theoretical known ratio and experimental area ratio are given. Injected CYP2E1 is calculated using the experimental ratio and injected aqua amount. Recoveries for CYP2E1 are given in % and are derived by comparing injected known amount and calculated amount.

RecombinantCYP2E1 loaded onto gel [pmol]	labelled peptide added to digest [pmol]	2E1 in-gel digest injected [pmol]	Labelled peptide injected [pmol]	Theoretical ratio native/labelled	N	Average calculated area ratio	stdev	sem	CV [%]	Calculated 2E1 amount [pmol]	2 E1 recovery [%l]	Accuracy ± 15 %
5	5	0.5	0.5	1	3	1.16	0.14	0.08	12	0.58	116	x
15	5	1	0.33	3	3	3.71	0.60	0.35	16	1.22	122	(x)
5	15	0.33	1	0.3	4	0.08	0.02	0.01	29	0.08	23	
5	25	0.2	1	0.2	3	0.07	0.01	0.01	20	0.07	36	
15	3	1	0.2	5	3	4.88	0.13	0.07	03	0.98	98	x

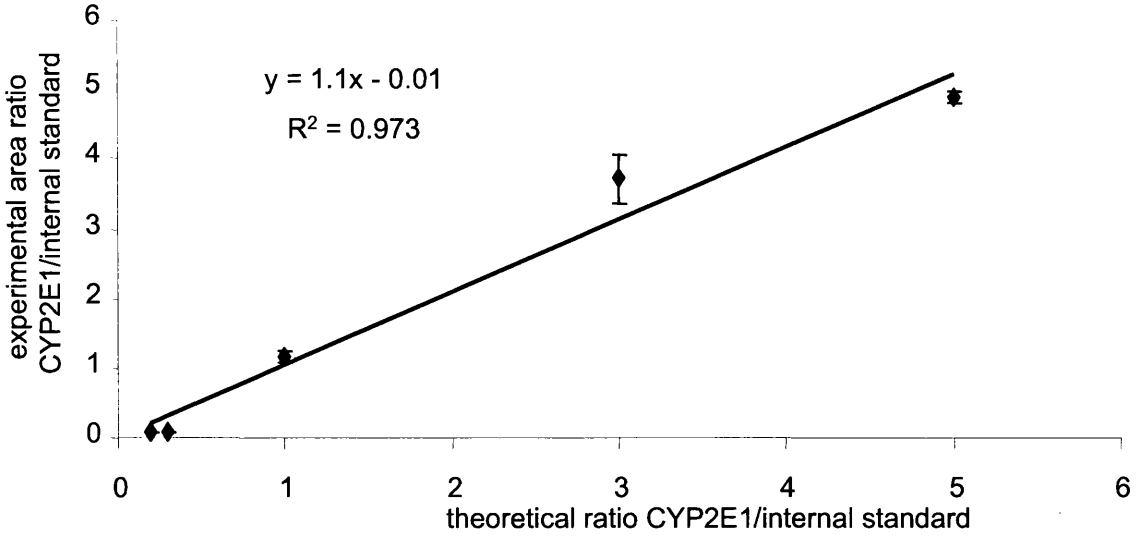


Figure 4.5: Graph for in-gel digestion calibration experiment: theoretical ratio (amounts of native to internal standard peptide) to experimentally detected area ratio. Error bars are given as standard error of the mean.

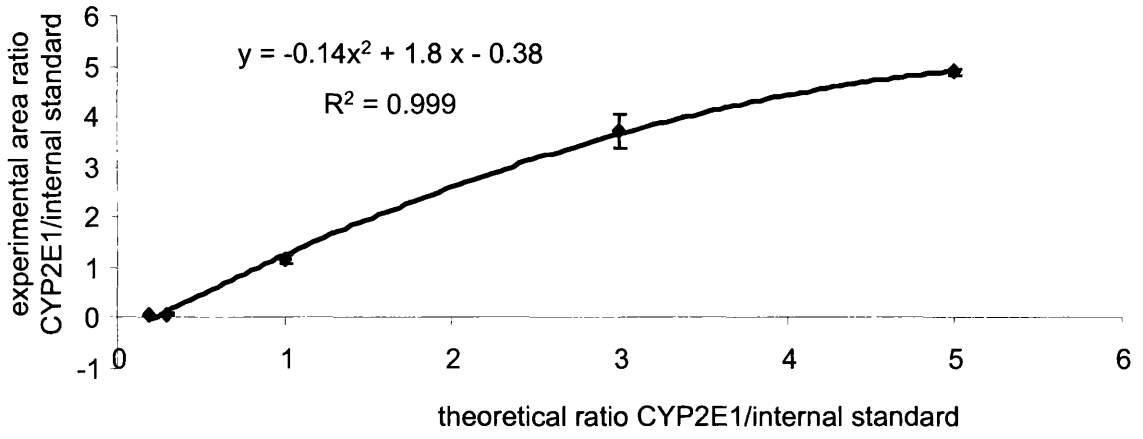


Figure 4.6: Applying quadratic equation to plot of theoretical versus experimental ratio in in-gel digestion experiment.

Table 4.5: Recalculation of protein content and recovery of calibration samples (Table 4.4) using the calibration line for in-gel digest (Figure 4.6, †$y=-0.14x^2 + 1.8x-0.38$).

CYP2E1 digest injected [pmol]	Labelled peptide injected	Theoretical ratio native/labelled [pmol]	N	Calculated area ratio (with equation†)	stdev	CV [%]	Recalculated 2E1 [pmol]	2 E1 recovery [%]
0.5	0.5	1	3	0.94	0.09	10	0.47	94
1	0.33	3	3	3.11	0.68	22	1.03	103
0.33	1	0.3	4	0.27	0.01	5	0.27	82
0.2	1	0.2	3	0.26	0.01	3	0.26	131
1	0.2	5	3	4.99	0.35	7	0.99	99

4.4.3 Spiking in biological sample (Standard addition)

In order to create a system as close to the experimental conditions as possible, microsomal liver samples were spiked with recombinant CYP2E1. The mixture was then treated as any other biological sample would be, including in-gel-digest and subsequent peptide extraction. Samples were then submitted to the same LC-MS/MS protocol as used for the in-solution and in-gel calibration experiments, monitoring the transitions as

described in 4.2. As the amount of recombinant protein added is known, the signal is a combined response of endogenous 2E1, abundant in the liver microsomes, together with the added recombinant protein. Both will show the same fragmentation channel 1281.8 → 1054.3. The amount of 2E1 determined using the internal standard represents the sum of both. Additions of 4 pmol, 8 pmol and 12 pmol were made to the same amount of liver microsomal protein (25 μg). In addition to the spiked samples, one sample without any recombinant protein addition was run as well.

Table 4.6: Spiking of recombinant CYP2E1 in human liver microsomal preparations. The total 2E1 amount is the sum of CYP2E1 in the sample and the amount spiked. A theoretical CYP2E1 amount is calculated as the sum of the amount spiked and amount detected in the unspiked sample. These values are compared with the total experimentally detected CYP2E1 to calculate the recovery rates.

Injected microsomal protein [g]	Injected recombinant CYP2E1 [pmol]	Injected labelled peptide [pmol]	N	Experimental area ratio native/labelled peptide	stdev	sem	CV [%]	total experimental CYP2E1 amount [pmol]	Calculated CYP2E1 amount in sample [pmol] (recombinant CYP2E1 +0.16)	Recovery (total experimental/calculated CYP2E1 [%]
2.5	0	0.4	10	0.41	0.09	0.03	22	0.16	0.16	
2.5	0.4	0.4	3	1.52	0.28	0.16	18	0.61	0.56	108
2.5	0.8	0.4	9	2.04	0.21	0.07	10	0.81	0.96	85
2.5	1.2	0.4	7	3.10	0.40	0.15	13	1.24	1.36	91

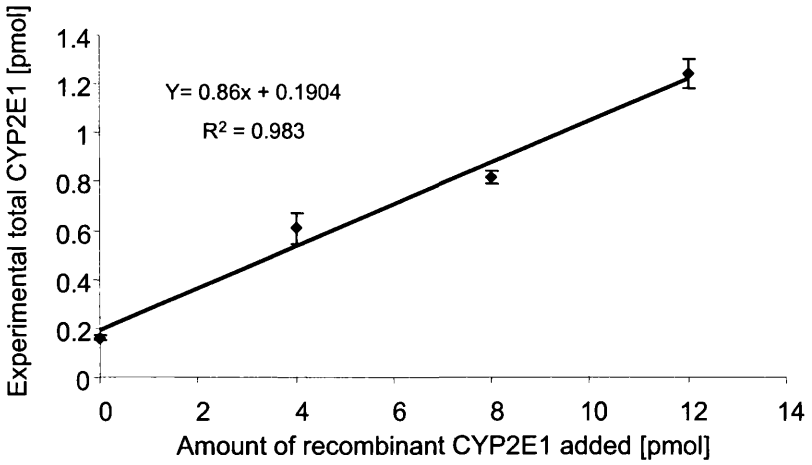


Figure 4.7: Spiking experiment of liver microsomal samples with recombinant CYP2E1. Plotted calculated total CYP2E1 amount versus experimentally measured total amount. Error bars are given as standard error of the means.

Table 4.6 shows the statistics for the spiking of recombinant CYP2E1 into liver microsomes and the quantification of these mixtures. Figure 4.7 shows the correlation between experimentally measured total amount of CYP2E1 and the theoretical values. There was a linear relationship between theoretical and experimental data. Using the area ratio of light to heavy labelled peptide the total 2E1 amount in the samples was calculated. The accuracy of this experiment could be measured by the agreement of theoretical and experimental values. Therefore the CYP2E1 value with no additional recombinant protein allowed the determination of CYP2E1 naturally abundant in this liver sample (0.16 pmol/25 µg microsomal protein injected). To determine the recovery as the ratio of experimental to theoretically expected CYP2E1 amount, the theoretical value was calculated by adding the spiked amounts (0.4, 0.8 and 1.2 pmol) to the determined CYP2E1 amount in liver microsomes (0.16 pmol). Comparing this to the experimental one, recoveries of 85%, 91% and 108% followed, meaning they are in a ± 15% interval from the theoretical value. The CV-values (10%–22 %) and standard errors demonstrated good/moderate precision of the data acquired in this experiment. The CVs for the higher amounts of CYP2E1 injected (0.96 pmol and 1.4 pmol) were well below the 15% required by the FDA[254], and only the CV for the lowest amount injected (0.16 pmol) was slightly above the CV required for the limit of quantification (22% *vs* 20%).

4.5 Method parameters

4.5.1 Limit of detection and quantification

The lower limit of quantification i.e. the lowest amount of CYP2E1 that could be quantified confidently, and the dynamic range were determined. In order to establish a lower limit of quantification, decreasing amounts of recombinant CYP2E1 were run on SDS-PAGE and digested with the same amount of labelled peptide (3 pmol). Previous experiments indicated that 1 pmol on gel is detectable, therefore amounts between 1 pmol and 0.25 pmol of recombinant CYP2E1 were loaded onto the gel and each digested with 3 pmol of heavy-labelled peptide. A chromatographic peak was only accepted for quantification if it showed a signal to noise ratio of 3 to 1 and minimum peak intensity of 10,000 cps. As described in the experimental section, one tenth of the

sample was injected per LC run, so injection amounts for recombinant CYP2E1 ranged between 0.1 and 0.025 pmol (1, 0.75, 0.5 and 0.25 pmol on gel). Only for the sample corresponding to 1 pmol on-gel the data fulfilled the requirements for acceptance. The lower amount injections resulted in peaks, which were outside the acceptance criteria. The lower limit of detection for the labelled-peptide was determined in a similar way. A fixed amount of recombinant protein (5 pmol) was in-gel digested with decreasing amounts of heavy labelled peptide (3, 1.5, 1, 0.75, 0.5 and 0.25 pmol). The limit for acceptance here was at 0.75 pmol labelled peptide. Therefore for further experiments in tissue samples, the lower limit of quantification was decided to be 1 pmol protein on gel and also for simplification 1 pmol of labelled peptide addition. This corresponds to a minimal injection of 100 fmol on the LC-MS/MS system (also equivalent to 40 pmol CYP2E1 per mg of microsomal protein or 200 pmol CYP2E1 per g of liver tissue). The highest amount of recombinant CYP2E1 loaded onto the gel was 15 pmol. As described above (4.4.2) there is a possibility that the ion trap becomes saturated when too much sample is injected. This value was then taken as an upper limit and defined that recombinant CYP2E1 is quantifiable by stable isotope dilution mass spectrometry over a range of 1–15 pmol on gel, (equivalent to 40–600 pmol/mg microsomal protein) when the isotope labelled peptide is at the 4–5 pmol level.

4.5.2 Selectivity

As biological samples are to be quantified after in-gel digestion there is a need to investigate whether the peaks detected and used for quantification originate uniquely from the CYP-derived peptide and its labelled analogue or anything else from the microsomes-gel matrix. Although MRM displays considerable sensitivity, selectivity is always an issue and any possible disturbing effect needs to be eliminated.

The first experiment was to analyse a piece of the gel, which was used for separation. It was subjected to the in-gel digestion protocol (without any protein or peptide addition) and run with the designed MRM method. Given that there was neither protein nor peptide added, no signal in the selected mass ranges should be detected. Indeed, the reconstructed ion chromatograms showed no distinctive peaks, with sufficient intensity or signal to noise, and also the MS/MS spectra in the known time-range do not derive from the appropriate peptide fragmentation (Figure 4.8). As

the gel matrix showed no interfering signals with the MRM method, the in-gel-digest of recombinant CYP2E1 (without labelled peptide) was analysed. The MRM transition appropriate for the heavy-isotope labelled peptide failed to generate a peak at the appropriate retention time and fragmentation of its precursor ion did not show a fragmentation pattern commensurate with the labelled peptide (Figure 4.9). A similar experiment was next performed with a liver sample (in-gel digested without labelled peptide) to investigate any interference of the biological matrix with the labelled peptide (Figure 4.10). The reconstructed ion chromatograms for the labelled peptide MRM did not show a signal, meeting the detection criteria (Figure 4.10 C) and the monitored fragmentation of *m/z* 1285.3 did not show the fragmentation pattern appropriate for the peptide (Figure 4.10 D). Thus the peaks in the *m/z* 1285.3 → 1057.6 reconstructed ion chromatogram did not derive from fragmentation of the heavy isotope labelled peptide parent ion. Figure 4.11 depicts the reconstructed ion chromatograms for the CYP2E1 MRM experiment applied to an empty gel piece submitted to in-gel digestion but with added 3 pmol of labelled peptide (0.3 pmol injected on-column). The MS/MS spectrum of the ion *m/z* 1281.8 eluting at the expected time for the light peptide is shown in panel D. The MRM detection requirements were not met, neither did the MS/MS spectrum correspond to that of the light peptide. This confirmed that in absence of the analyte detectable peaks were absent from the appropriate MRM chromatograms. The controls carried out proved the selectivity of the MRM method for the chosen parent masses. Thus it was decided that the quantification of CYP2E1 could be carried out in microsomal preparations confidently.

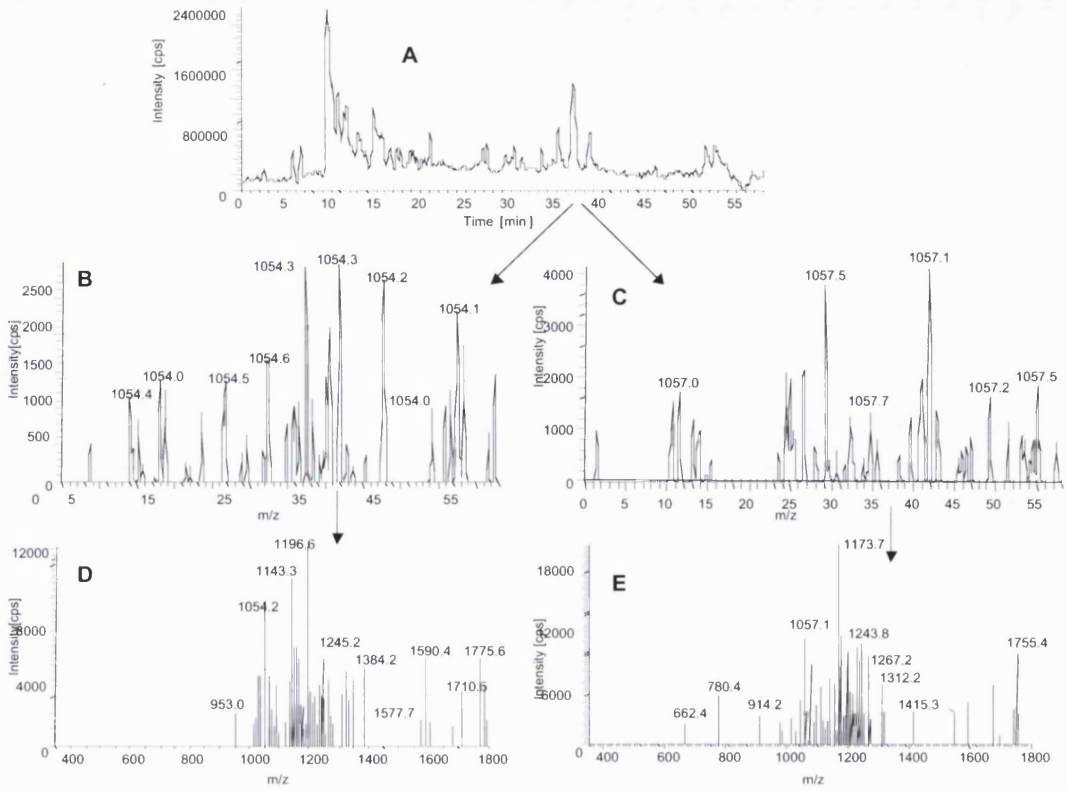


Figure 4.8: Interference test of gel-matrix: A: chromatogram for whole LC-MS run. B: reconstructed ion chromatogram for monitoring the light peptide (*m/z* 1281.8 → 1054.3) and (C) for monitoring the labelled internal standard peptide (*m/z* 1285.3 → 1057.6). D: MS/MS spectrum of the precursor ion *m/z* 1281.8, and E of the precursor ion *m/z* 1285, both eluting at about 37 min (arrows indicate elution time)

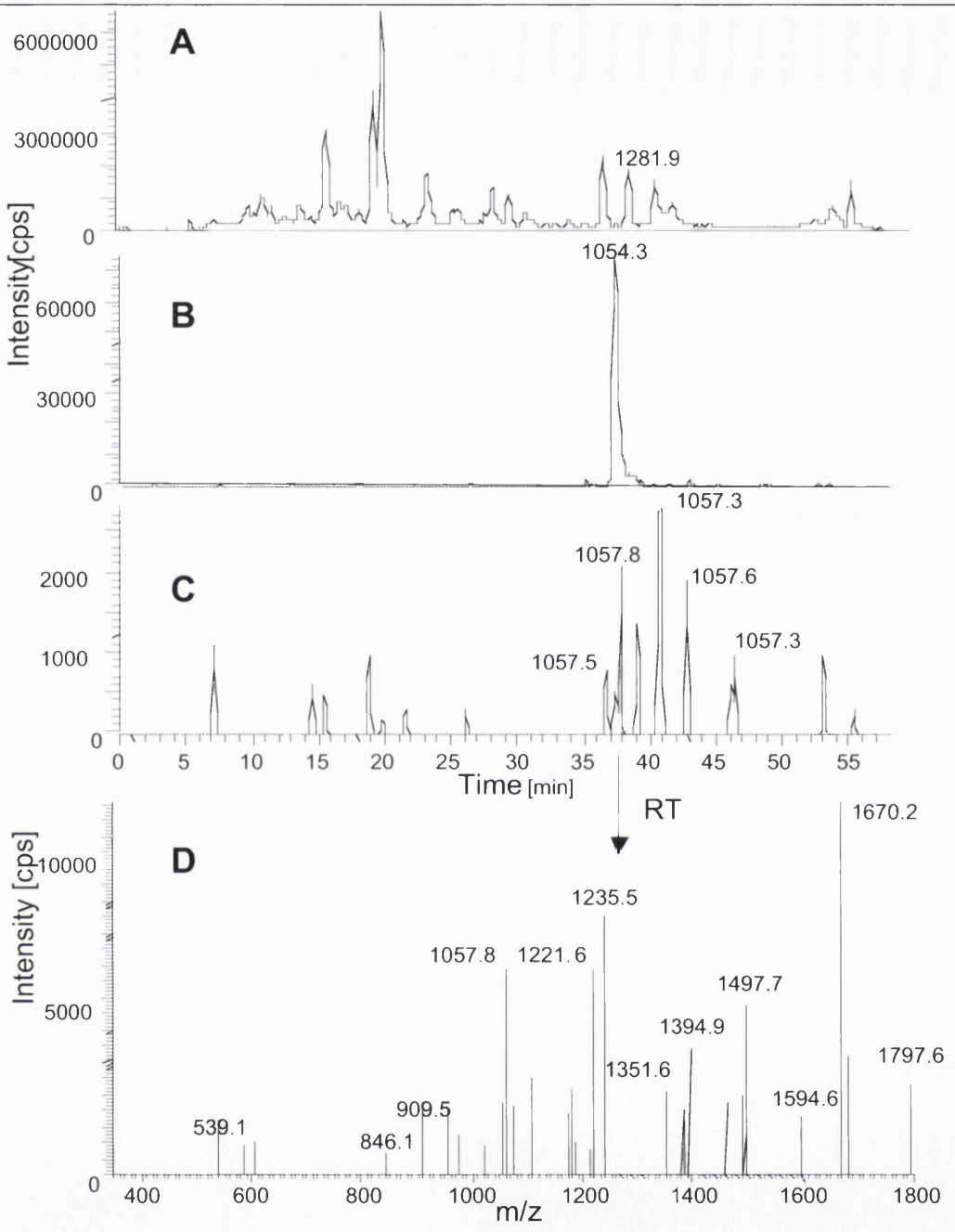


Figure 4.9: Injection of recombinant CYP2E1 in-gel digest. A: Total ion chromatogram for complete run; B: Reconstructed ion chromatogram for monitoring the light peptide transition *m/z* 1281.8 → 1054.3 and (C) for monitoring the labelled internal standard peptide transition *m/z* 1285.3 → 1057.6. D: MS/MS spectrum of the ion of *m/z* 1285.3 eluting at theoretically expected elution time of 37.5 min observed for the native peptide (1281.8).

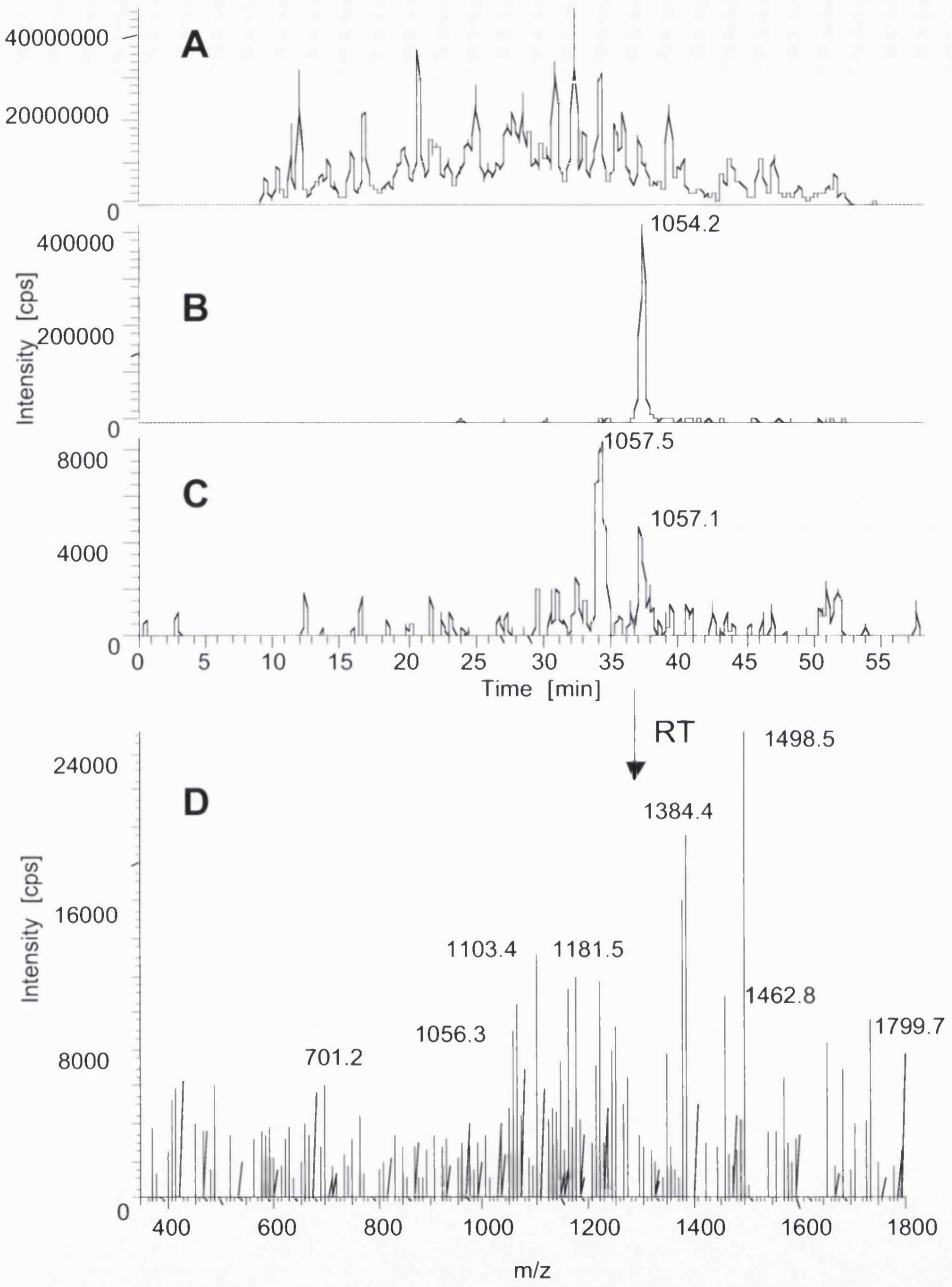


Figure 4.10: Analysis of liver microsomes using the CYP2E1 MRM method A: chromatogram of LC-MS-run. B: reconstructed ion chromatogram for monitoring of light peptide (*m/z* 1281.8 → 1054.3). C: reconstructed ion chromatogram for monitoring of heavy peptide (*m/z* 1285.3 → 1057.6). D: MS/MS spectrum for fragmentation of heavy parent ion *m/z* 1285.3 eluting at about 36.8 min as expected from the elution of the native peptide (1281.8).

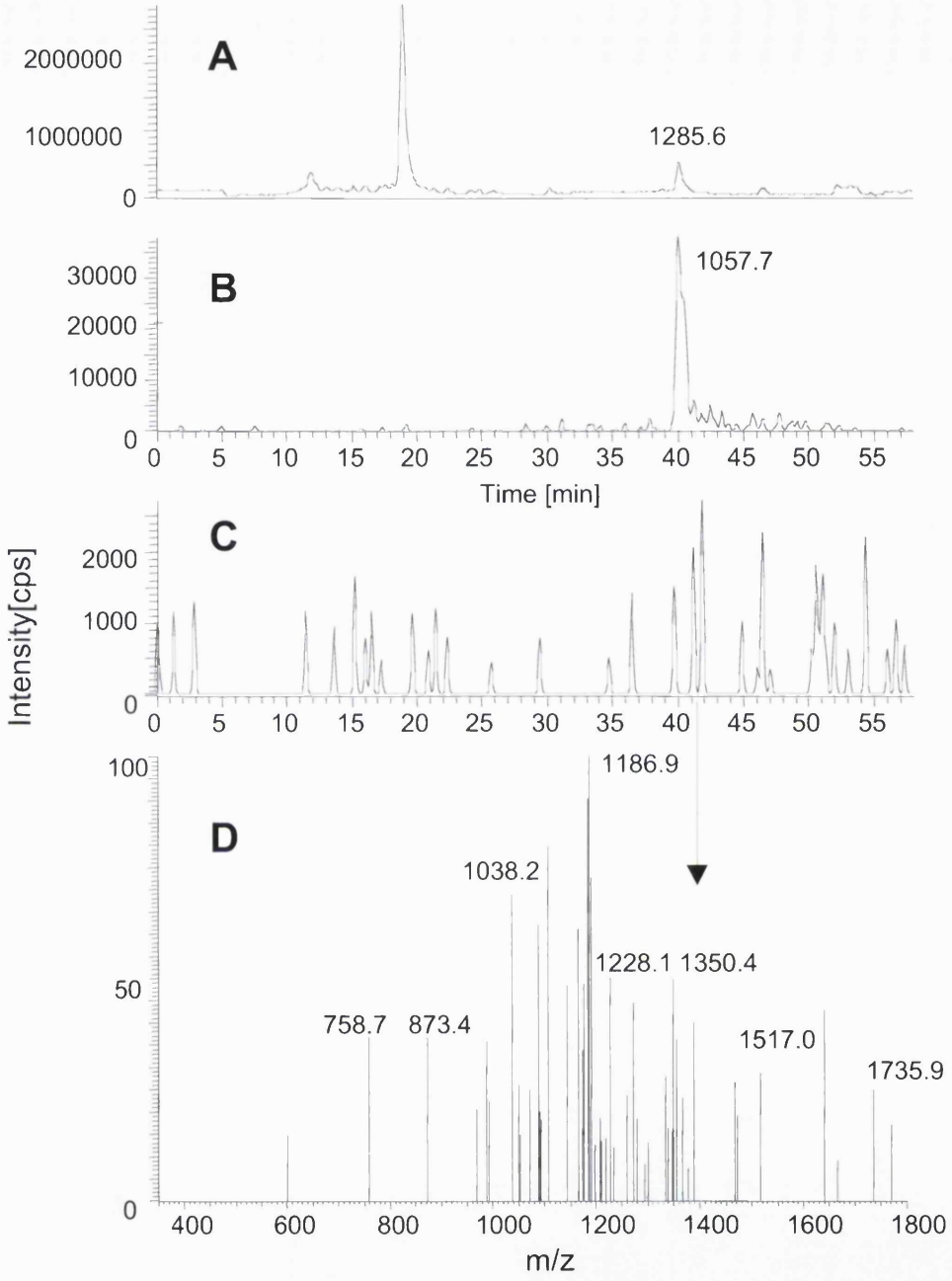


Figure 4.11: MRM analysis of an empty gel piece in-gel digested with 3 pmol labelled peptide addition. A: chromatogram of LC-MS-run. B: reconstructed ion chromatogram for monitoring of isotope labelled internal standard peptide (*m/z* 1285.3 → 1057.6) C: reconstructed ion chromatogram for monitoring of light peptide (*m/z* 1281.8 → 1054.3). D: MS/MS spectrum for fragmentation of light parent ion m/z 1281.8 did not result in the expected fragment eluting at about 40 min.

4.6 Quantification of CYP2E1 in human liver microsomes

4.6.1 Individuals and protein content

Human liver tissue samples from four individuals were analysed. Liver pieces derived from surgeries for liver tumour or colon metastases were processed using the protocol described in 2.2.1. Prior to gel electrophoresis it was necessary to determine the total protein content in the samples. Bradford analysis was always performed on the day the sample was run on the gel, so protein-content variation due to freeze-thaw cycles could be eliminated. This was an important aspect as the CYP2E1 quantity was referred to the total microsomal protein content. The Bradford assay was carried out in duplicate for liver samples as well as for the BSA-standards. A typical calibration curve is shown in Figure 4.12. Due to individual variation of concentration of the microsomes, the sample had to be diluted prior to the assay. Usually dilutions of 1:50, 1:100, 1:500, and 1:1000 were tested to achieve absorbance values within the standard curve. The absorbance measured for a microsomal sample was used to calculate the total protein content.

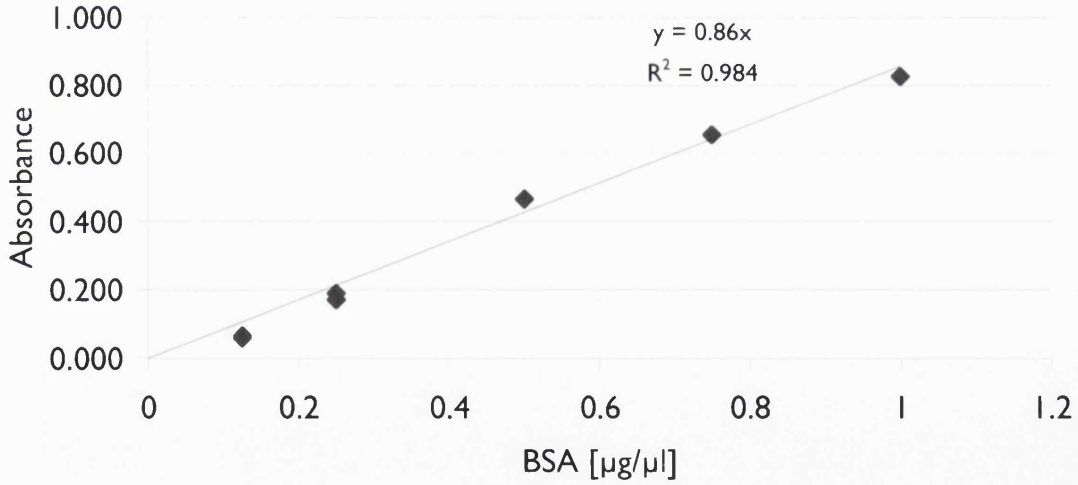


Figure 4.12: Calibration line for total protein determination using the Bradford assay. The calibration line was created with bovine serum albumin (measured in duplicate). The graph equation was used for calculation of total protein content in liver microsomes.

Table 4.7 gives an overview of the total content of all proteins detectable by Coommassie G-250 in each sample and the microsomal protein amount applied on the gel. In general 25 µg of microsomal protein was loaded onto the gel, as this was the maximal protein amount appropriate for the gel-system used. Although the total protein content in tumour microsomes was lower when compared with normal tissue, the low number of samples does not allow to determine whether this difference is statistically significant with a sufficient power; using a paired *t*-test, no statistically significant difference was found ($p>0.05$).

Table 4.7: Total protein amount (determined via Bradford assay) and amount of microsomal protein used for gel electrophoresis. (HN= "liver normal" or control, HT= "liver tumour")

Sample	Mass of tissue used [g]	Total volume of microsomes produced [ml]	Total protein content [µg/µl]	Protein amount loaded onto gel [µg]
H1N	0.63	2	6.09	24.97
H1T	0.66	0.3	3.4	23.8
H2N	0.5	0.4	8.85	26.55
H2T	0.5	0.25	3.2	25.6
H3N	0.64	0.75	4.72	23.6
H3T	0.72	0.5	3.08	24.64
H4N	0.7	0.75	3.98	23.88
H4T	0.63	0.5	2.4	25.2

4.6.2 Gel electrophoresis and in-gel digestion

Microsomal samples were separated (for protein amounts see Table 4.7) by 1D-SDS PAGE and gel pieces were selected for in-gel digestion. As the molecular weight of CYPs ranges between approximately 45–62 kDa, the corresponding area in the gel was excised and prepared for in-gel digestion. A generous area in this mass range was cut out and divided into 5 bands with one bigger central band around 50 kDa and two smaller bands above and below. Figure 4.13 shows a typical gel of liver microsomes and the mass range cut out is indicated. Recombinant CYP2E1 was run as well as an indicator of the correct mass range (Figure 4.13, lane 5).

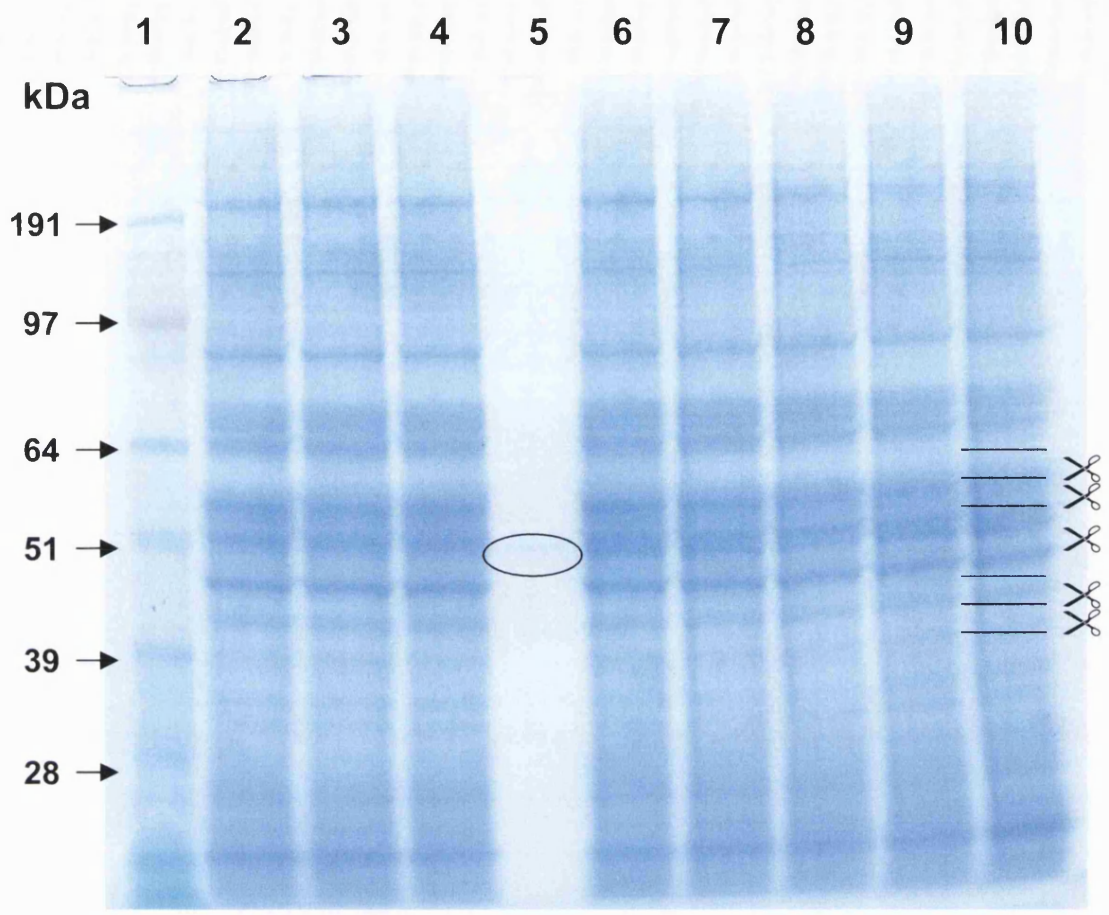


Figure 4.13: 1D-SDS-PAGE of microsomal liver preparations. Lane 1: Molecular weight marker. Lane 2 to 4 and 6 to 10: liver microsomes from individual H1N. Lane 5: recombinant CYP2E1, 5 pmol applied (highlighted spot). Selected bands for cutting are indicated on the right hand side with one bigger central band and two below and two above. All lanes were cut out accordingly.

4.6.3 Protein identification

Each band was digested, extracted and first submitted to MS/MS analysis in order to identify proteins and show the presence or absence of CYPs in the prepared bands. The microsomal preparations of the five individuals were subjected to database search using the Mascot search engine and the CYPs identified in these bands documented. Table 4.8 lists the CYPs found in the analysis. CYP2E1 was found in all individuals amongst other CYPs such as CYP1A2, CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP3A4, CYP4A11 or CYP4A22. CYP2E1 was only found in the central 50 kDa band so only

this band was analysed further. This can be seen in Figure 4.14, which shows the example of subject H2N run with the CYP2E1 MRM-method.

Table 4.8: Identified CYP enzymes in subjects H1N to H4N. Proteins are listed in descending hit-order. Peptides with peptide scores > 35 were accepted. Uniqueness of peptide was proofed via Blast search (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) performed against a non-redundant human protein database. Information on protein identification in these liver samples can be found on the accompanying CD.

Sample name	Protein accession number	Protein name	Protein score	Protein sequence coverage [%]	Number of peptides identified	Number of unique peptides
H1N	IPI00007282	CYP2A6	225	23.5	9	3
	IPI00299568	**CYP2E1**	187	22.7	8	8
	IPI00307246	CYP1A2	125	7.4	3	3
	IPI00019160	CYP4F2	101	10.4	4	0
	IPI00007219	CYP2C9	74	6.7	2	1
	IPI00013323	CYP2C19	69	4.9	2	1
	IPI00429358	CYP4A22	68	2.3	1	0
	IPI00290301	CYP2C8	57	2.7	1	1
	IPI00793171	CYP3A5	45	9.9	1	1
	IPI00010218	CYP20A1	42	3.2	1	0
	IPI00103355	CYP2F1	39	2.6	1	1
H2N	IPI00007219	CYP2C8	141	6.7	3	1
	IPI00290301	CYP2C9	134	6.7	2	2
	IPI00003480	CYP2A13	126	10.3	4	0
	IPI00299568	CYP2A6	123	13	5	1
	IPI00007282	**CYP2E1**	53	4.7	1	1
	IPI00164895	CYP3A7	47	2.1	1	0
	IPI00429358	CYP4A22	46	2.3	1	0
	IPI00103355	CYP2F1	44	2.6	1	1
H3N	IPI00465138	CYP3A4	425	30.8	11	4
	IPI00007219	CYP2C9	392	13.7	6	2
	IPI00299568	CYP2A6	318	24.3	9	2
	IPI00307246	CYP1A2	307	14	5	5
	IPI00290301	CYP2C8	293	10.8	5	4
	IPI00302839	CYP2A6	198	14	5	2
	IPI00007282	**CYP2E1**	194	20.1	8	7
	IPI00025831	CYP3A5	165	10.8	5	2
	IPI00465089	CYP4A11	113	4	2	1
	IPI00657852	CYP2C19	74	8.1	3	0
	IPI00002370	CYP4F3	59	2.7	2	0
	IPI00019411	CYP2J2	47	2.6	1	1
H4N	IPI00007282	**CYP2E1**	72	4.7	1	1
	IPI00003480	CYP2A13	69	2.6	1	0
	IPI00465138	CYP3A4	54	4.6	2	0
	IPI00007219	CYP2C9	40	2	1	0
	IPI00793171	CYP3A5	38	9.9	1	1

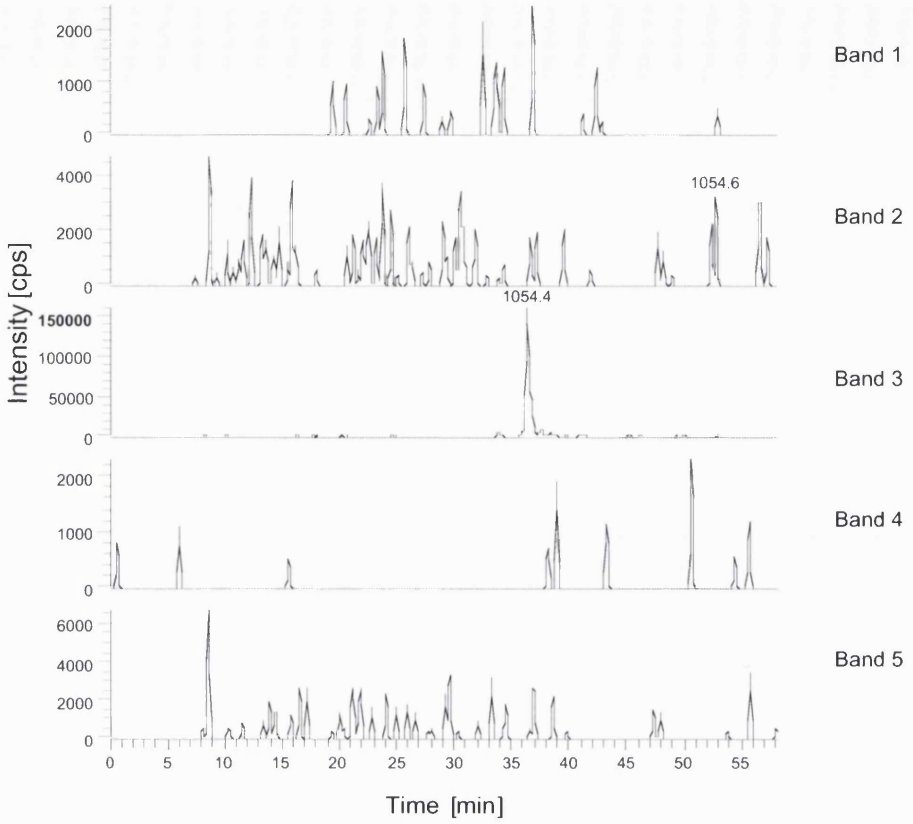


Figure 4.14: Reconstructed ion chromatogram for the CYP2E1 MRM run (m/z 1281.8 → 1054.3) of all five gel bands of subject H2N. The monitoring of the fragmentation of the light peptide shows a high response only for band 3, the central band.

4.6.4 Intra-sample run clean up

In all quantification experiments it is important to reduce carry-over effects to a minimum. As the hydrophobic character (4.2) of the peptide was known, an organic solvent was chosen to wash the peptide completely off the column before a new sample was injected. In fact, after trials with different numbers of blanks it was decided that only after three blank injections it could be assured that there was no carry over from a previous run.

4.6.5 CYP2E1 Quantification

For quantification of CYP2E1 in human liver microsomes the samples were in-gel digested in presence of heavy labelled internal standard peptide and after extraction

subjected to MRM analysis. For each sample three lanes were run on the gel and each band was injected in triplicate, this resulted in (a maximal of) nine injections per sample. However the raw data had to fulfill certain criteria in order to be accepted. The signal to noise ratio had to be at least 3 to 1 and a minimal peak intensity of 10,000 cps was necessary, so data from less than nine replicates was used for further analysis. The areas under the peaks were experimentally determined and as the amount of internal standard peptide was known, the quantity of endogenous CYP2E1 peptide was calculated, and the endogenous protein content inferred. The CYP2E1 content is given relative to the amount of total microsomal protein (88–200 pmol CYP2E1 per mg microsomal protein) as well as in relation to the weight of the wet tissue extracted (0.38–2.51 nmol/g wet tissue). The quadratic equation ($y = -0.14x^2 + 1.8x - 0.38$) determined in the in-gel digestion calibration experiment (4.4.2) was used to calculate CYP2E1 quantities as well as the linear equation ($y = 1.1x - 0.01$). The results for quadratic calculation are labelled in red in Table 4.9. The difference between values calculated using the linear and quadratic equations is less than 20%. For smaller quantities, such as subject H4N, the two values are very similar. As expected, the higher the CYP2E1 content, the bigger the difference between the two values. Because of the better correlation coefficient for the quadratic calibration line, this function provides precision. It should be noted that the inter individual variation in CYP2E1 levels is considerably greater than the variation due to equation selection.

Table 4.9: CYP2E1 content (± standard deviation) from four individual samples of human liver microsomes calculated using the linear (top) and quadratic (bottom, red; $y=-0.14x^2 + 1.8x-0.38$ from Figure 4.6)) calibration function. The number of accepted runs for quantitative analysis is given as N.

Subject	N	CYP2E1 per injection [pmol]	CYP2E1 per mg microsomal protein [pmol]	CYP2E1 per g wet tissue [nmol]	CYP2E1 per g wet tissue [µg]	CV [%]
H1N	5	0.32 ± 0.01	129.76 ± 2.81	2.51 ± 0.05	137.1 ± 2.97	2.2
		0.29 ± 0.005	114.24 ± 1.79	2.21 ± 0.03	120.26 ±1.89	1.6
H2N	8	0.39 ± 0.11	144.85 ± 41.63	1.03 ± 0.29	55.96 ± 16.1	28.7
		0.33 ± 0.07	122.52 ± 27.21	0.87 ± 0.19	47.34 ± 10.51	22.2
H3N	8	0.47 ± 0.12	200.51 ± 51.99	1.11 ± 0.29	60.52 ± 15.69	25.9
		0.38 ± 0.08	162.66 ± 34.92	0.90 ± 0.19	49.10 ± 10.54	21.5
H4N	8	0.21 ± 0.05	88.12 ± 22.65	0.38 ± 0.1	20.49 ± 5.27	25.7
		0.21 ± 0.03	89.81 ± 13.93	0.38 ± 0.06	20.89 ± 3.24	15.5

The precision of the quantification in biological samples could be explored by looking into inter- and intra-batch variability. As only one microsomal preparation was processed per individual, the experiments of one preparation had to be compared. The same amount of microsomal protein from a given sample was loaded onto three parallel lanes of the gel. The CYP containing band was digested and the digest was injected three times. As mentioned earlier, the maximum of nine injections per subject was not achieved for all individuals (4.6.5), so calculated inter and intra batch coefficients of variation (CV) had to be based in some cases on fewer replicates. The triple injection of one band was treated as one batch. The variation between the results of these three injections could be seen as intra-batch difference and was given as coefficient of variation. It also gave an example of the repeatability of the method. The CVs for intra-batch variation ranged between 3% and 23%. The inter-batch variation was determined by comparing the average of the triple injection of each band, and indicated differences between the three bands of the same sample. The CVs for this ranged between 2 and 29 % and are given in Table 4.9. The CVs for the inter-batch evaluation were higher than for the intra-batch calculation.

4.6.6 Quantification in CYPs in liver microsomes derived from tumour tissue

The liver tumour tissue samples of subjects H1N to H4N were also processed and analysed (Figure 4.15). The MS/MS experiment as described for the liver samples in 4.6.3 were performed and no CYPs could be determined in these tumour microsomes. For confirmation of the identification result, MRM was performed on these samples using the CYP2E1 method and no CYP2E1 was determined.

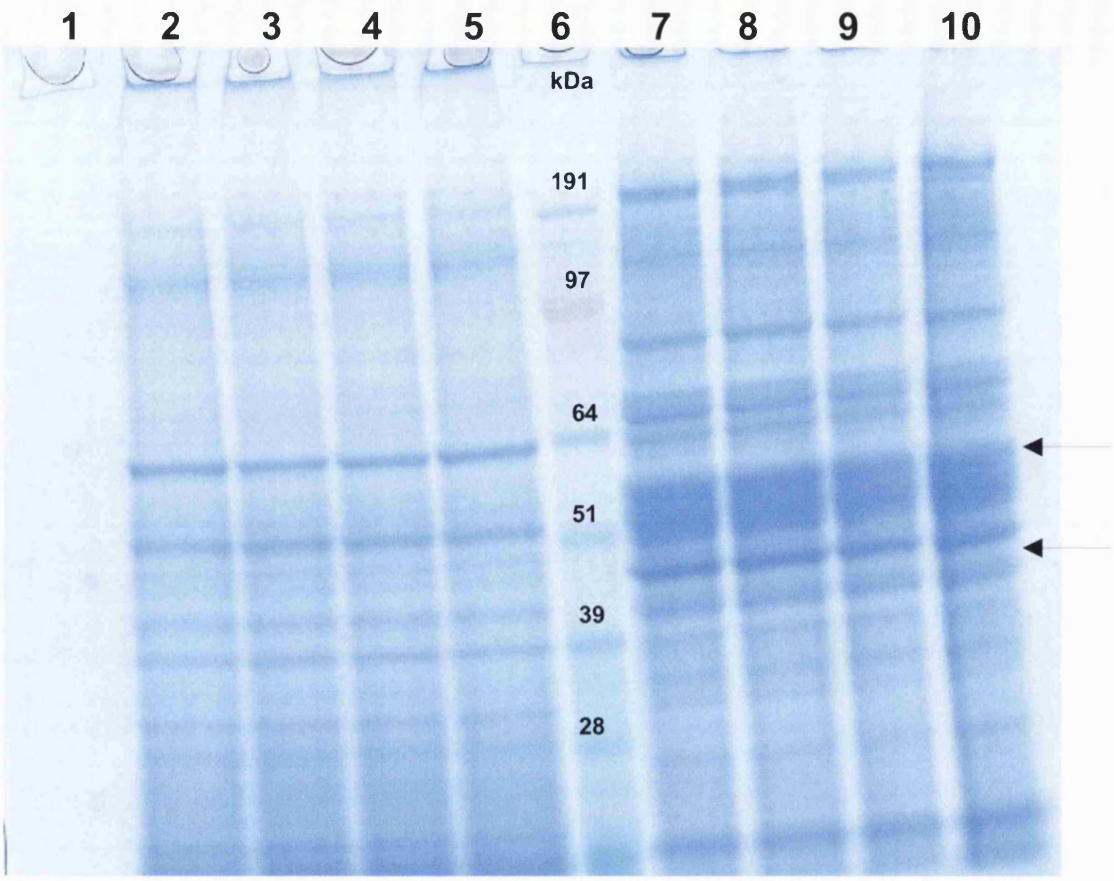


Figure 4.15: 1D-SDS-PAGE of microsomal liver preparartions of subject 2L. Lane 6: Molecular weight marker, lane 2 to 5: microsomes derived from tumour tissue. Lane 7 to 10: microsomes derived from control tissue. Lane 1 left empty. Arrows indicate CYP450 molecular weight range.

4.6.7 Quantification results in relation to data in literature

The range of CYPs identified by database searching agreed with previous data obtained using similar methodology [243], in other mass spectrometric studies [255] and by immunoblotting [216, 217]. The data apparent in literature regarding quantification of CYP2E1 in liver using diverse techniques led to lower but still comparable results. In a published immunochemical study, Shimada *et al* [217] showed a CYP2E1 content of 22 pmol/mg human liver microsomal protein (in our work we found 88 to 200 pmol CYP2E1/mg microsomal protein). They determined the total CYP concentration in microsomes to be ~350 pmol/mg microsomal protein. Afsar *et al.*, [256] found the total CYP content of rat liver microsomes to be 1 nmol/mg protein. In an earlier immunochemical study

Rodrigues [216] determined the CYP2E1 content of human liver microsomes to be 49 pmol/mg protein. In conclusion, and considering that CYP2E1 represents about 10% of total microsomal CYP protein [216, 217] the data presented in our study is in good agreement with published values. For further comparison, using mass spectrometry and a modified ICAT method, the level of CYP2E1 in mice liver microsomes was determined to be 35.13 pmol/mg total protein [257].

It is note worthy that the CVs of 24% to 30% determined by the method employed here are of the same order as those determined by Barnidge *et al.* [155] in their original publication describing a method to quantify proteins by using tryptic peptides and stable isotope dilution mass spectrometry

4.7 Resume and Discussion

The developed method for quantification of CYP2E1 described herein via multiple reaction monitoring using a stable isotope labelled internal standard was successful in quantifying the protein in microsomal liver samples. This is the first attempt to apply this technique to quantify CYP2E1 in human liver microsomes. However, validation of the method was necessary in order to design a reliable and robust method, which could function eventually as a routine-method. From the choice of the right parent ion and its fragments for selective reaction monitoring to the study of accuracy, precision and sensitivity, all these factors were investigated before the method was approved. The direct infusion of the peptide showed already that it was necessary for the peptide to be dissolved in a more organic solvent than is usual for tryptic peptides (4.2) and also the carry over in the blanks (4.6.4) showed that this peptide exhibited a hydrophobic character. This had an influence on the behavior of the peptide (and the recombinant CYP2E1 protein) in the LC as it was a late eluting compound. Additionally, this no doubt contributed to the underestimation of CYP2E1 in in-solution digests, as the labelled peptide may not have been fully solubilised. The fact that a longer gradient was used than is usual for tryptic peptide analysis would make the method less suitable for "high throughput", so gradient adaptation i. e. shorter total run time would be desirable. This would also be an issue if CYP2E1 had to be quantified together with rather early eluting CYPs in one sample, however the introduction of UPLC or monolithic columns may circumvent this problem. The difference in recovery from in-

solution and in-gel digest (4.4.1, 4.4.2) suggests that the internal standard peptide was less soluble in the in-solution digestion buffer and therefore the recoveries were very high (between 172 and 246 %). When the peptide is added to the in-gel digestion sample, it is more easily solubilised and hence, the recoveries are more accurate than compared to the in-solution digest experiment. The tendency seen for the in-gel calibration curve to follow a quadratic rather than a linear relation (Figure 4.6) showed that overloading of the ion trap might have an influence on the quantification result and thus a determination of lower and upper limit of quantification was necessary, as was a suitable quantitative addition of labelled peptide. The calculation of CYP2E1 quantification results in liver microsomes using the quadratic equation showed the biggest deviation for the sample containing the highest amount of CYP2E1 of 19%. The lower the amount of CYP2E1 present, the closer the values derived from using the quadratic equation or the area ratio for content calculation. The best imitation of the experimental conditions was provided with the spiking experiment, which showed acceptable results regarding recoveries as well as precision (4.4.3). Finally, the well recognized benefits of the MRM method, namely sensitivity and selectivity for special fragments, was demonstrated and no interference from sample or matrix could be detected (4.5.2). The advantage of combining this technique with the already well established in-gel-digestion protocol exhibited at the same time a big challenge as many factors could have an influence on the outcome of the quantification results. Therefore validation should not be underestimated, as only this way weaknesses of the method can be detected.

The CYP2E1 concentrations determined here in microsomal protein where higher than those published elsewhere [216, 217, 257] but still in a similar range and the differences can be attributed to differences in methodology, sampling procedure and sample processing. The amount of CYP2E1 found in liver tumour microsomes was 141 ± 23 pmol/mg microsomal protein with an inter-individual variation between 88 pmol/mg and 201 pmol/mg, resulting in a coefficient of variation of 33%. This is lower than the inter-individual variation found by Shimada *et al* [217] (55%) even though this study investigated CYP2E1 in samples from 60 subjects.

Chapter 5 Quantification of Cytochrome P450 1A2

5.1 Introduction

The successful quantification of CYP2E1 described in chapter 4 lead to the development of the stable isotope dilution method for the quantification of other CYPs. Taking into consideration the most abundant CYPs already known in human liver, such as 1A2, 2A6, 2B6, 2C8, 2C9, 2C19, 2D6, 2E1 and 3A4, a second one, CYP1A2 was chosen for quantification, because it was found to be present in sufficient quantities in the samples available. Its abundance in human liver is about 13% of total liver CYPs [217] and it is also involved in the metabolism of xenobiotics in the liver [258]. CYP1A2 plays an important role in the activation of carcinogenic compounds such as arylamines and nitrosamines, both present in tobacco [223]. The information of the expression state of CYP1A2 in a particular tissue can therefore provide useful information of its susceptibility to toxic damage caused by metabolic activation via CYP1A2. Although the quantification method for CYP2E1 was validated before, the new CYP required once more a thorough validation of the method prior to analysis of tissue samples.

5.2 Selection of internal standard peptide for CYP1A2

Initially, a suitable stable isotope labelled peptide for CYP1A2 had to be selected. The peptide with the sequence IGSTPVLVLSR was chosen as the best candidate for CYP1A2 because of its mass spectrometry compatible length (11 amino acids) and the presence of an arginine residue at the *C*-terminus. After the solubility issues experienced with the long CYP2E1 peptide (23 amino acids), it was decided to choose a smaller peptide in order to avoid similar problems e.g. with the LC-gradient. This peptide was also selected because it had been identified previously on the same instrument using a similar method of data dependent MS/MS detection.

5.3 Fragmentation pattern and MRM method design

In contrast to CYP2E1, the CYP1A2-peptide was diluted with *only* 0.1 % formic acid to obtain stock and working solutions, as it showed good solubility in this solvent. Validation was performed as described for CYP2E1 (chapter 4). Initially, the detection potential in MS and MS/MS modes was determined. A 100 fmol/µl solution of the labelled peptide was infused on its own. The labelled CYP1A2 peptide (IGSTPVL*VLSR) showed clearly (Figure 5.1) the doubly charged peak ($[M+2H]^{2+}$ m/z 575.3†). The triply charged peak ($[M+3H]^{3+}$ m/z 383.8) was not visible in the spectrum. The fragmentation patterns were next investigated by direct infusion of an equimolar mixture of recombinant protein digest and the corresponding isotope labelled peptide. Figure 5.2 shows the fragmentation pattern of the native (A) and labelled peptide (B) specific to CYP1A2 and in both spectra the most intense signal was that for the y_7 - fragment-ion (cleavage *N*-terminal to proline), which was used in the MRM method.

During the acquisition of the fragmentation patterns of the labelled peptide, the instrument was tuned on this peptide separately. An equimolar mixture of isotope-labelled peptide and recombinant protein digest was used to acquire MS as well as MS/MS spectra using the different instrument parameters. It was found that the CYP1A2-peptide tune parameters were not resulting in better peptide identification than using the standard instrument settings for [Glu^1]-Fibrinopeptide B (GluFib) $[M+2H]^{2+}$ – a standard used for proteomics experiments. Thus, it was decided to use the established Glufib-tune parameters for the MRM method to quantify CYP1A2. In the first scan event the fragmentation of the light peptide of CYP1A2 was monitored (m/z 571.8→783.5), in the second scan event, the heavy labelled peptide fragmentation was observed (m/z 575.3 → 790.5).

† Unless stated otherwise, all masses in this chapter are the average mass.

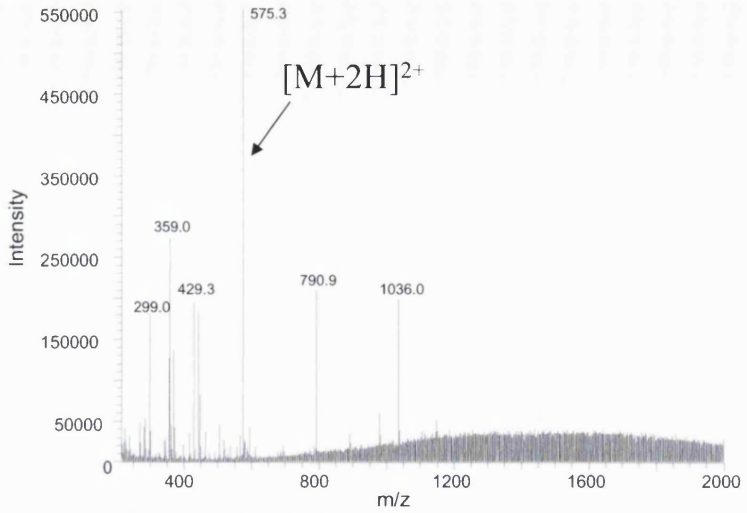


Figure 5.1: Direct infusion (100 fmol/µl) of heavy isotope labelled CYP1A2-specific peptide IGSTPVL*VLS R. Doubly charged peak visible (here *m/z* 575.3)

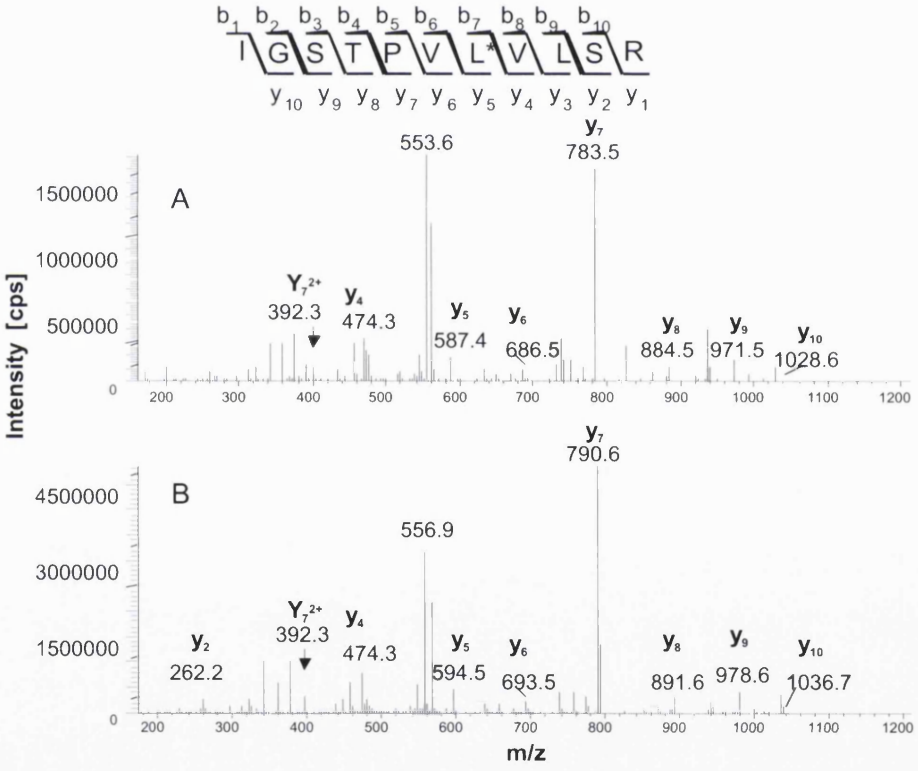


Figure 5.2: Infusion of recombinant CYP1A2 in-solution digest with CYP1A2 internal standard peptide: Fragmentation pattern of CYP1A2-specific (A) unlabelled (IGSTPVLVLSR) ($[M+2H]^{2+}$ *m/z* 571.7) and (B) stable isotope labelled peptide (IGSTPVL*VLSR) $(M+2H]^{2+}$ *m/z* 575.3). The ions at *m/z* 553.6 and 556.9 correspond to $[M+2H-2(H_2O)]^{2+}$ ions. Water loss is expected from the S and T residues.

Table 5.1: Overview of *m/z* values used for monitoring of CYP1A2-fragmentation

Protein	Peptide	Mass of peptide (average)	theoretical *m/z* of $[M+2H]^{2+}$ ion (average *m/z*)	observed *m/z*	Z of observed parent ion	observed *m/z* of fragment ion for MRM
CYP1A2	IGSTPVLVLSR	1141.38	571.70	571.8	2	783.5 [y_7]
	IGSTPVL*VLSR	1148.38	575.19	575.3	2	790.5 [y_7]

5.4 LC-method development

The selected peptide for CYP1A2 eluted earlier than those selected for CYP2E1, so a much shorter gradient of 38 minutes could be used. Figure 5.3 shows the elution of the two pairs of CYP-specific peptides at two different time points for the injection of the peptide/recombinant protein digest mixture. The transitions defined in Table 5.1 were applied and the reconstructed ion chromatograms were generated. In Figure 5.3 the chromatograms A and B depict the transition for the CYP1A2 specific light peptide and heavy labelled peptide respectively. Both elute at around 17.5 minutes.

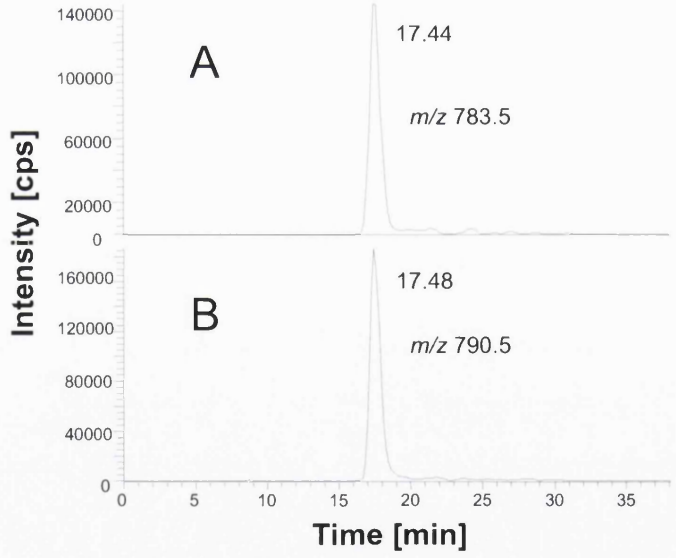


Figure 5.3: Reconstructed ion chromatograms of all transitions in the CYP1A2 MRM experiment. For the native CYP1A2-specific peptide (IGSTPVLVLSR) the transition 571.8 → 783.5 (A) and for the heavy peptide (IGSTPVLVL*SR) 575.3 → 790.5 (B) were monitored.

5.5 Selectivity of MRM method

The experiments depicted in Figure 5.3 showed that the applied LC-MS-MRM method can separate the CYP1A2-specific peptides and selectively monitor the transitions into specific product ions. However, this has been shown only with authentic standards and in-solution digests of recombinant protein ignoring the possible influence of the gel matrix or the liver sample itself. The gel matrix provided no interference to the detection method (Figure 5.4). There were no intense signals in the expected time range (approximately 17 minutes), and no peptide-specific fragmentation spectra could be identified (data not shown). The injection of an in-gel digest of liver microsomes without any internal standard peptide addition is presented in Figure 5.5. The transitions for the labelled peptide specific for CYP1A2 (C) could not be detected at the expected retention time and only the fragmentation of the native peptide (A) was evident, showing the absence of any interfering compounds. As there are no interfering entities this also illustrates the selectivity of the method to separate between light and heavy peptides.

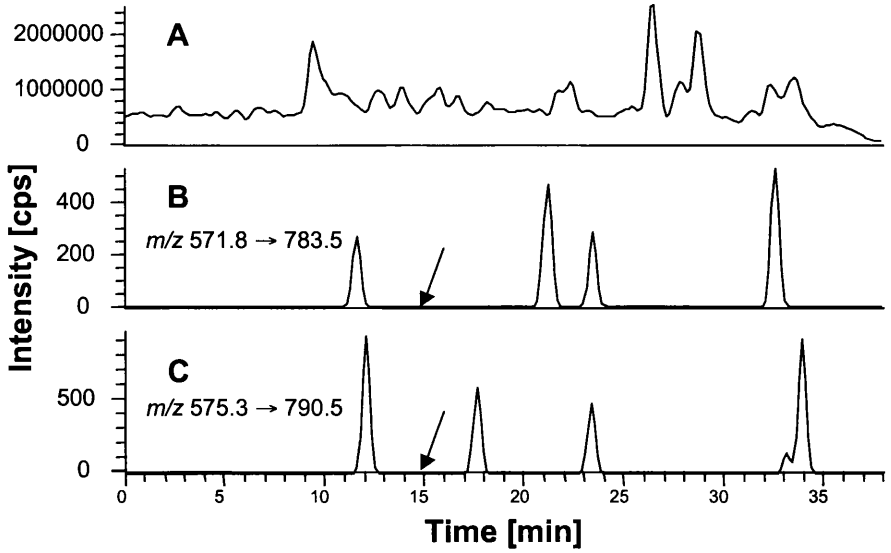


Figure 5.4: Reconstructed ion chromatograms for injection of a gel piece without sample or internal standard peptide addition. (A) Total ion chromatogram for LC-MS run. (B & C) The arrows indicate when elution of the respective peptides would be expected in the respective chromatograms. The arrows indicate the elution-time of CYP1A2-specific peptides (IGSTPVLVLSR and IGSTPVL*VLSR).

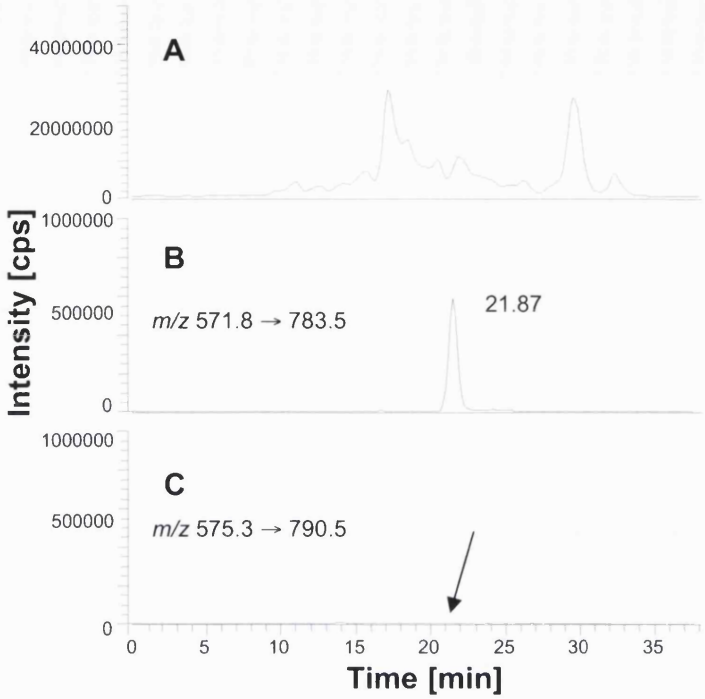


Figure 5.5: Reconstructed ion chromatograms of liver microsomal injection after in-gel digestion without internal standard peptide addition. (A) total ion chromatogram for LC-MS run. (B & C) Transitions selected for light and heavy CYP1A2-specific peptides. The arrow indicates when elution of the heavy peptide would be expected in the respective chromatogram. Heavy CYP1A2-specific peptide (IGSTPVL*VLSR) does not elute at the expected 21.8 minutes.

5.6 In-solution digest calibration

In-solution digests were prepared as explained in 2.2.9. Different amounts of recombinant protein digest of CYP1A2 were mixed with the same amount of labelled peptide (always 2 pmol) and injected using the LC-MS-MRM method for CYP1A2. The experimental ratios calculated from areas of native and internal standard peptide were plotted versus the theoretical known ratios. Thus, in each case 200 fmol of internal standard peptide were injected together with in-solution digests of recombinant CYP1A2 ranging between 10 and 800 fmol. In this way more ratios, i.e. more data points for the calibration curve were available compared to CYP2E1 calibration (chapter 4). Additional to the transitions defined in 5.3, the formation of other product ions was monitored and the results compared.

5.6.1 CYP1A2 in-solution calibration

The fragmentations of the precursor ions (*m/z* 571.8 and 575.3) into the most intense fragments (*m/z* 783.5 and 790.5) were monitored, as were the second most intense fragment ions (*m/z* 553.6 and 556.9). Reconstructed ion chromatograms for the following additional transitions were generated and areas compared: *m/z* 571.8 → 553.6 and 575.3 → 556.9. Injections of the lower amounts of recombinant CYP1A2 peptides (10 fmol and 20 fmol) resulted in data which did not meet the acceptance criteria (minimal signal intensity of 10,000 cps and signal to noise ratio of at least 3 to 1). For the most intense transitions data was accepted from 40 fmol injections onwards. In the case of the second pair of product ion transition (*m/z* 571.8 → 553.6 and 575.3 → 556.9) data was accepted at a level of 60 fmol CYP1A2 digest injection. The calibration lines are shown in Figure 5.6 (Table 5.2 for data) for both pairs of transitions monitored. The linear regression is slightly better for the more intense fragment transition (R^2= 0.956) and also the CV values are better. Monitoring the transition into *m/z* 783.5 and 790.5 fragments also showed a higher sensitivity as injections of lower peptide amounts lead to acceptable peaks and therefore this fragmentation channel was used for quantification. The observed difference between theoretical and experimental ratio can be explained by low solubility of recombinant protein in the digestion buffer, remembering that CYP 1A2 is a membrane bound protein.

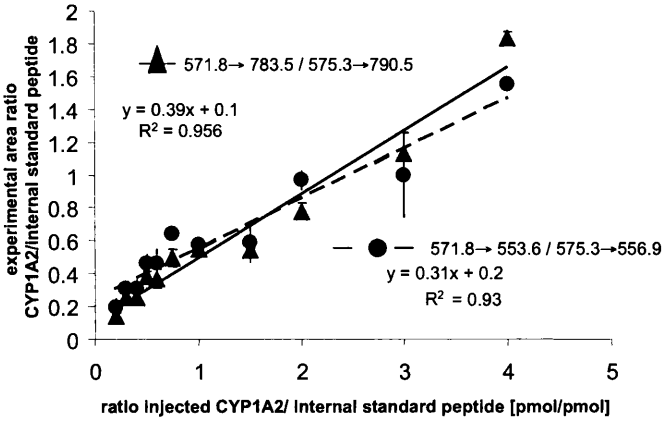


Figure 5.6: Calibration using an in-solution digest of recombinant CYP1A2. Calibration lines for the two pairs of transitions are plotted. Error bars are given as standard error of the mean. The graph shows the ratio of the amounts injected and the corresponding peak area ratios

Table 5.2: Comparison of LC-MS/MS analysis of CYP1A2 in-solution digest using different amounts of peptide and internal standard and using different transitions (*m/z* 571.8 → 783.5 and 571.8 → 553.6 as well as 575.3 → 790.5 and 575.3 → 556.9). The lower experimental ratio is presumably due to the lower solubility of the intact protein in the digestion buffer (isp: internal standard peptide; N number of accepted injections).

Amount peptide injected [fmol]			Monitored transition *m/z* 571.8→783.5 and 575.3→790.5					Monitored transition *m/z* 571.8→553.6 and 575.3→556.9				
CYP1A2	Internal Standard Peptide (isp)	CYP1A2/isp [pmol/pmol]	N	Peak area ratio ratio (CYP1A2/isp)	Standard Deviation	sem.	CV [%]	N	Peak area ratio ratio (CYP1A2/isp)	Standard Deviation	sem	CV [%]
40	200	0.2	3	0.14	0.03	0.01	17.44	2	0.20	0.001	0.001	0.91
60	200	0.3	3	0.25	0.02	0.01	7.44	3	0.31	0.08	0.05	27.43
80	200	0.4	5	0.26	0.06	0.03	21.88	5	0.31	0.04	0.02	13.74
100	200	0.5	3	0.38	0.06	0.03	15.83	3	0.46	0.11	0.06	23.67
120	200	0.6	4	0.37	0.11	0.06	30.15	4	0.46	0.11	0.05	23.12
150	200	0.75	5	0.50	0.12	0.05	23.2	5	0.64	0.20	0.09	31.10
200	200	1	3	0.55	0.06	0.04	11.81	3	0.58	0.03	0.02	5.5
300	200	1.5	3	0.54	0.02	0.01	3.14	3	0.59	0.06	0.03	9.89
400	200	2	3	0.78	0.09	0.05	11.48	3	0.97	0.21	0.12	21.92
600	200	3	2	1.13	0.17	0.12	15.44	2	1.0	0.08	0.06	8.06
800	200	4	2	1.83	0.06	0.04	3.19	2	1.55	0.36	0.26	23.42

5.7 CYP1A2 in-gel digestion calibration

The in-solution calibration represented rather artificial conditions for digestion, whereas a system incorporating the gel matrix and the associated protocol had more bearing on the exact quantitative methodology to be used. Therefore recombinant CYP1A2 was separated on a 1-D gel, in-gel digested and peptides extracted; these samples were then submitted to LC-MS using the CYP1A2 LC-MS-MRM-method. Amounts of recombinant protein loaded onto the gel varied between 750 fmol and 10 pmol. Standard additions of 2 pmol of internal standard peptide were made. The reconstructed ion chromatograms for both pairs of transitions were monitored; however, only the transitions of *m/z* 571.8 → 783.5 and *m/z* 575.3 → 790.5 showed a signal to noise ratio of at least 3 to 1 and a minimal signal intensity of 10,000 cps. Data was only accepted for the injection of 300 fmol of recombinant in-gel digests and more. The lower limit of quantification loaded on the gel was thus 3 pmol, (which is equal to 300 fmol injection into the

LC-MS-system). Upper limit of quantification was 10 pmol applied on the gel or 1 pmol injected on-column, higher amounts were not investigated (Table 5.3). However, the injection of large amounts of protein can lead to non-reproducible results, due to column and ion-trap overload effects. The relationship between theoretical and experimental ratios was approximated using a linear and a second order polynomial function. The correlation coefficient for the second-order polynomial function with R^2=0.988 was better than for a linear regression, using the points for lower CYP1A2 injection only (R^2=0.93) (Figure 5.7). The data showed that the polynomial function represents the tendency of the ratios to be low or very close to the expected ratio at lower proportions of recombinant CYP1A2 digest, but increasing disproportionately for higher amounts injected to more than the double than expected. Higher area ratios were found for the injection of higher concentrated samples. A possible explanation for this is the suppression of ionisation of the internal standard peptide by high amounts of analyte and solubility effects. This was clearly observed at a ratio of more than 3.5 to 1 (native/internal standard peptide, Table 5.3). Only few points of the calibration curve therefore seem to be located in the linear range of this experiment, corresponding to a range between 300 and 700 fmol of native peptide injected or 3 to 7 pmol recombinant protein added onto the gel. This has to be considered when determining quantities in biological samples. For further calibration experiments it would be useful to use more points in the linear range.

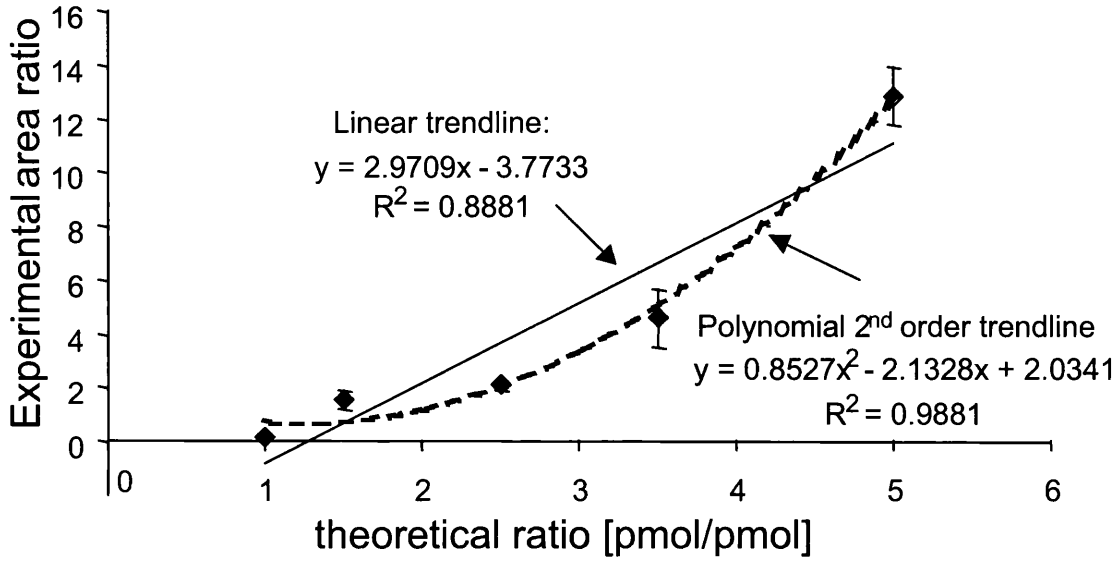


Figure 5.7 Calibration using an in-gel digest of recombinant CYP1A2. Error bars are given as standard error of the mean. The graph shows the ratio of the amounts injected and the corresponding peak area ratios. A possible polynomial 2nd order trend-line is indicated with a dotted line. In the range between 1 and 3.5 of theoretical ratio, a linear equation was used.

Table 5.3: LC-MS/MS analysis of recombinant CYP1A2 in-gel digestion using different amounts of peptide and internal standard. (isp: internal standard peptide; N: number of accepted injections)

Amount peptide injected [fmol]			Monitored transition 878.7→860				
CYP1A2	**Internal Standard Peptide (isp)**	**CYP1A2/isp [pmol/pmol]**	**N**	**Peak area ratio (CYP1A2/isp)**	**Standard Deviation**	**sem**	**CV [%]**
200	200	1	1	0.17	—	—	—
300	200	1.5	3	1.57	0.61	0.35	38.3
500	200	2.5	3	2.08	0.27	0.16	13.1
700	200	3.5	3	4.59	1.86	1.08	40.62
1000	200	5	3	12.84	1.87	1.08	14.45

5.8 Bradford assay of recombinant CYP1A2

Due to the difficulties experienced with the calibration experiments, such as low recoveries of native peptide especially from the in-solution digestion experiments and the possibility that the properties of the protein could be responsible for insufficient recoveries, the recombinant protein was analysed using a Bradford assay. The supplier guaranteed a purity of > 90 % for recombinant CYP1A2. The total protein quantity determined by the Bradford assay of the protein was to be compared to the concentration declared by the producer: Bradford analysis was performed as described in 2.2.2. It should be noted that the Bradford assay determines the total protein content, so if there were any other protein-contaminations in the recombinant protein preparation sample, they would be detected as well. As the protein was expressed in *Escherichia coli* (*E. coli*), the presence of peptides from the host organism in the sample was also investigated. Using a database search without restrictions to human proteins, it was possible to identify several proteins from *E. coli*. However, this is not surprising as the Certificate of Analysis guarantees only 90% purity for 1A2.

For the CYP1A2 protein content determination, there was 93% of the specified protein detected. These results were not unreasonable considering the lack of specificity of the assay, additionally, a difference of 7% could be explained by sample degradation, as proteins had been stored for two years. A more accurate determination of protein content could be achieved by quantitative amino acid analysis, but this has not been performed here.

Table 5.4: Bradford assay of recombinant proteins. The recovery is calculated as the ratio between the determined concentration and the nominal concentration according to the supplier.

Recombinant protein	Bradford assay [mg/ml]	Supplier [mg/ml]	Recovery [%l]	CV [%l]
CYP1A2	1.198	1.29	93	4

5.9 Digestion trials for digestion efficiency of CYP1A2

The in-solution calibration experiments showed low recoveries for the native peptide. A possible explanation is the incomplete digestion of the protein into the peptide, which had to be further investigated. During digestion with trypsin, different factors can

influence the digestion efficiency and had to be investigated. The use of trypsin has the advantage that it is a widely used enzyme in protein digestion experiments and cleavage sites are well known [87]. It also had to be distinguished between in-solution and in-gel digestion, of which the latter includes more experimental steps introducing possible sample losses such as gel cutting or digest extraction from the gel.

5.9.1 In-solution trials

The trypsin to protein-ratio usually used in proteomics experiments is between 1:100 and 1:20 (as also recommended by the manufacturer). The in solution digestion of 5 µg recombinant CYP1A2 was investigated by varying enzyme: protein ratios; 1:50 as a low trypsin to protein proportion, 1 to 30 and 1 to 10 as very high trypsin to protein ratios. Additionally to the assessments of the use of different trypsin amounts, it was also tested whether a further trypsin addition after four hours of digestion would enhance digestion. All digestion trial samples were injected and analysed by the LC-MS-MRM-method and signals in the reconstructed ion chromatograms for native and internal standard peptide transitions were compared. After in-solution digestion, an equimolar amount of CYP1A2 labelled peptide was added and a sample containing 100 fmol of this peptide was injected. Six samples differing in trypsin amount and number of trypsin additions were investigated in this way (Table 5.5). The areas of the native peptide were observed as a measurement of increase in digested peptide. Samples are named by the trypsin: CYP1A2 ratios: 1:50, 1:30 and 1:10. A single trypsin application is labelled with "+" and "++" indicated a second addition of the same amount of trypsin after 4 hours. Except for the 1:30++ and the 1:10+ samples the cv values were acceptable (between 1 to 14 %). If the low trypsin amount was sufficient for complete digestion, one would expect no significant differences in the areas derived from digestions using a higher amount of trypsin. Using a one-way ANOVA, no overall significant difference between areas could be detected ($p>0.05$, Figure 5.8).

Table 5.5: Comparison of in-solution digests of CYP1A2 with different amounts of trypsin. Overnight (14h) digests were started with different amounts of trypsin, corresponding the trypsin/protein ratio shown. "++" indicates a second addition of the same amount of trypsin after four hours incubation. For LC-MS/MS analysis, internal standard peptide (same amount as native CYP1A2) was added.

Trypsin/Protein ratio	Total Trypsin amount added [ng]	Average peak area (CYP1A2)	Standard Deviation	%CV
1:50 +	100	3.85E+06	4.92E+05	12.8
1:50 ++	200	3.88E+06	1.79E+05	4.6
1:30 +	167	4.78E+06	1.31E+05	2.7
1:30 ++	334	5.08E+06	1.03E+06	20.3
1:10 +	500	3.63E+06	9.19E+05	25.3
1:10 ++	1000	4.11E+06	1.92E+05	4.7

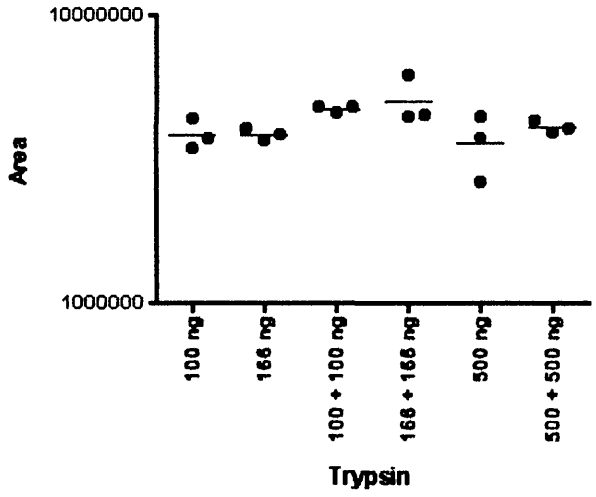


Figure 5.8: Comparison of different amounts of trypsin in in-solution digests of CYP1A2. An additional amount of trypsin was added to some samples after four hours. For LC-MS/MS analysis, internal standard peptide was added and the resulting peak area was plotted versus the total amount of trypsin used.

During the in-solution trial several different (but equimolar) concentrations of CYP1A2 protein/internal standard peptide-mixtures were injected. When all the injections were listed and the areas and ratios were compared the following results were found: The areas increased as the injected amount of CYP1A2 (in-solution digest) increased, as did

the area of standard peptide; but to a different extent. A change in internal standard peptide concentration lead to much stronger signal variation than for the native peptide, indeed the slope for the internal standard peptide was 20 times higher. This can be explained by a better solubilisation of internal standard peptide than recombinant protein in the buffer. As the recombinant protein is not soluble in the buffer, even injection of higher amounts do not result in the same increase in signal as observed for the internal standard peptide (Figure 5.9).

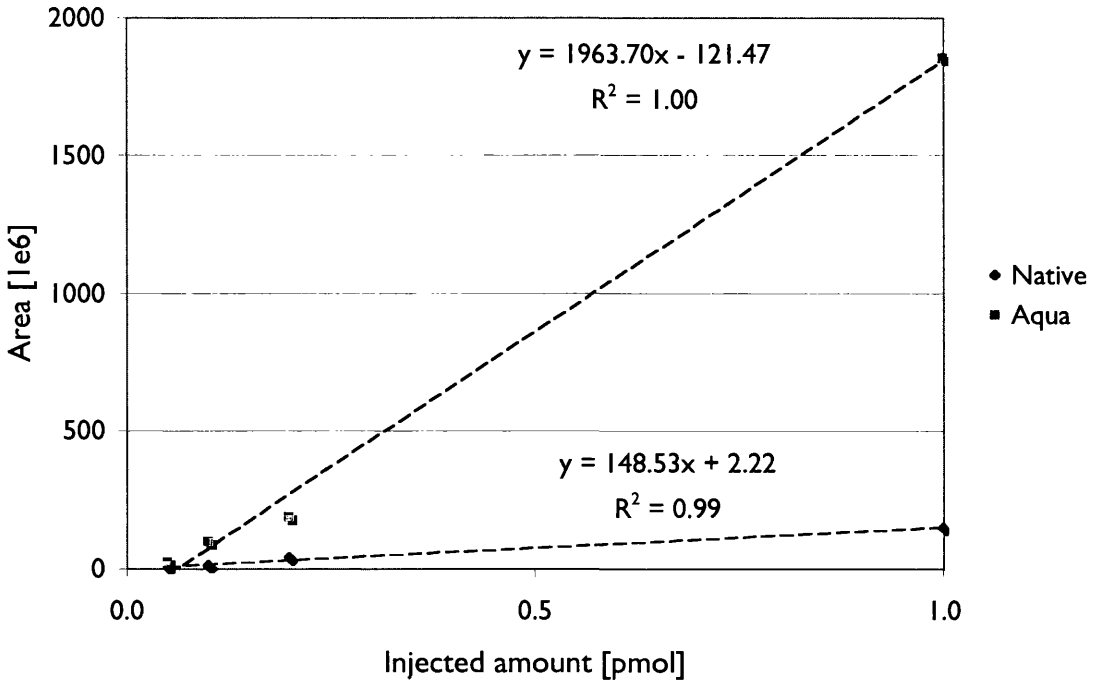


Figure 5.9: Relationship between the injected amount of sample of in-solution experiments of CYP1A2 and peak area of native peptide and internal standard peptide. Same amounts of native peptide and internal standard peptide were analysed using LC/MS/MS. The graph shows a different slope for the relationship between injected samples and peak area for native and internal standard peptide.

5.9.2 In-gel digestion trial

The digestion in the gel matrix and the extraction out of it is a very critical step regarding sample losses and interference with detection (for selectivity see 5.5). To simplify the experimental settings (in contrast to the in-solution trial), the amount of trypsin used for digestion was investigated only. In-house protocols [200] used for CYP-protein identification have previously used a minimum of 250 ng of trypsin per gel band. To compensate for a larger protein amount in biological samples and to ensure complete digestion, the trypsin amount in quantification assays was increased to 375 ng per band and more for bigger bands, depending on the protein content. Using recombinant protein digests from gel pieces, the effect of different trypsin amounts was evaluated. For all samples 5 pmol recombinant CYP1A2 were used, which equalled approximately 290 ng of protein (Figure 5.10). To an equimolar mixture of recombinant

protein and internal standard peptide 375, 750 and 1000 ng trypsin were added. Digestion was carried out as described in 2.2.8.4. Also here, the peak areas of native CYP1A2 peptide were plotted versus the trypsin amount added for in-gel digest. A one-way ANOVA was performed and no overall significant difference ($p>0.05$) was found among the experiments (Figure 5.11).

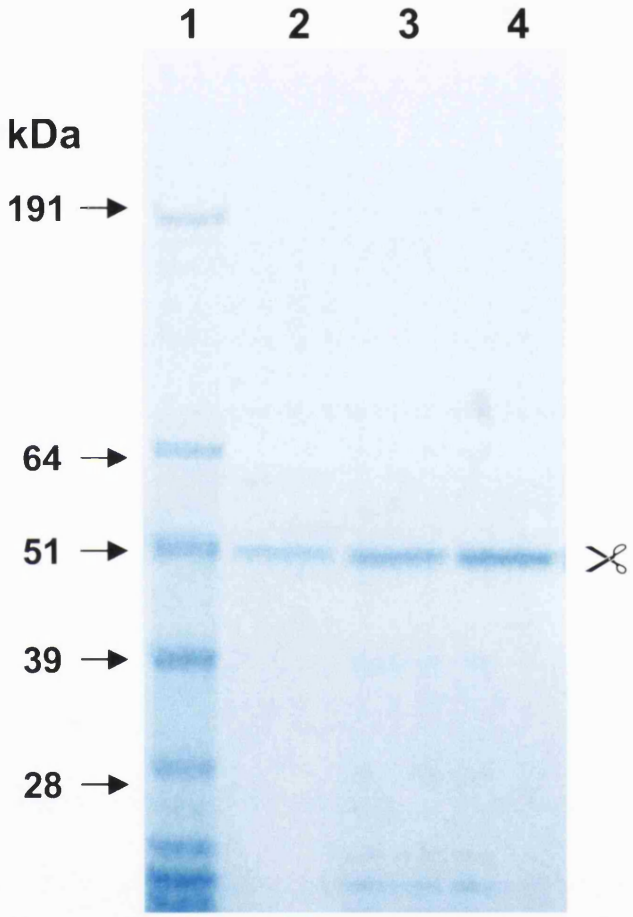


Figure 5.10: Typical SDS-PAGE of recombinant CYP1A2: Lane one contains molecular weight marker. In lane two to four 5 pmol of recombinant CYP1A2 are loaded. The scissors indicate the position were gel was excised.

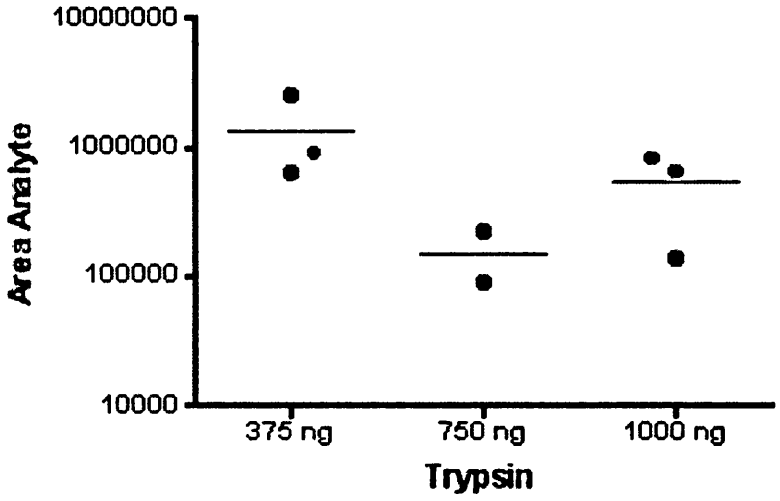


Figure 5.11: Comparison of areas for different amounts of trypsin in in-gel digests of CYP1A2. For LC-MS/MS analysis, internal standard peptide was added and the resulting peak area was plotted versus the trpyisn amount added.

As a further test for completeness of digestion, an in-solution digest of recombinant CYP1A2 was run on a gel. On the same gel in addition to the molecular weight maker undigested recombinant CYP1A2 was run as well (Figure 5.12). Assuming complete digestion, no band should be seen in the lane of the digest around the 50 kDa weight marker, which could be shown with this gel. However, it should be noted that the LC-MS method is far more sensitive than a Coomassie stain, which only allows a detection of around 7 ng of protein according to the manufacturer specification, and about 1 pmol in our experience.

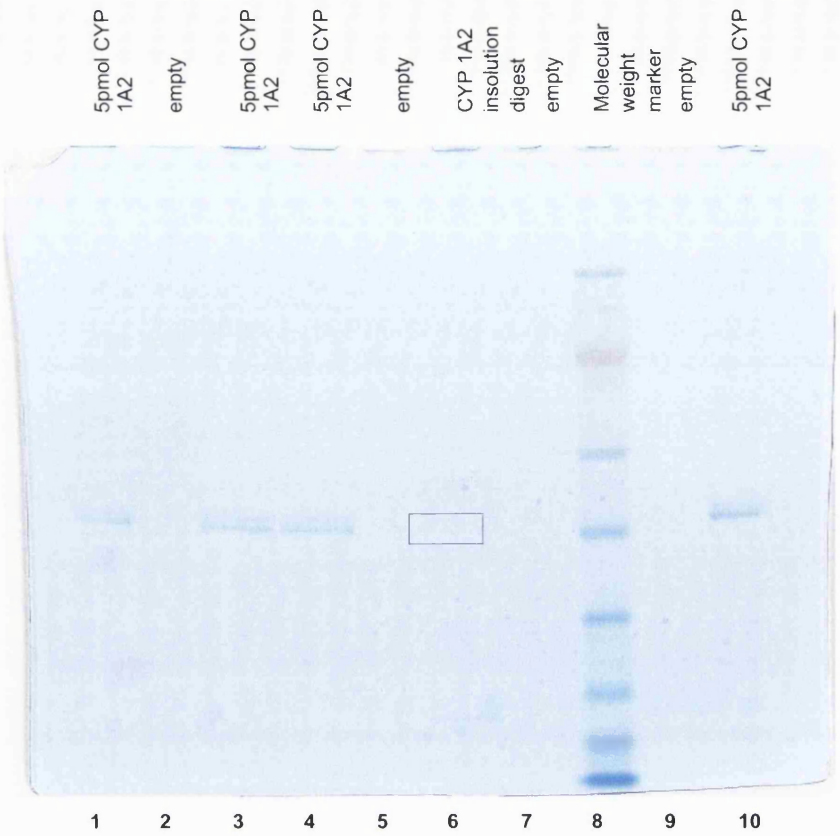


Figure 5.12: SDS-PAGE of recombinant CYP1A2 and in-solution digests of CYP1A2. Lanes 1, 3, 4 and 10 were loaded with recombinant CYP1A2, lane 6 contained in-solution digest of 5 pmol CYP1A2; lanes 2, 5, 7 and 9 were left empty. In lane 6 on the same height as the CYP1A2 bands in lane 1, 3, 4 or 10 there is no parallel band visible.

5.10 Analysis on QTOF

In order to validate the results, which were obtained during calibration of CYP1A2, some samples were run on a second instrument in order to compare and confirm the results. In-solution digest of CYP1A2 was mixed in three different ratios with internal standard peptide and injected. Because of the higher sensitivity of the QTOF not more than 200 fmol of protein digest was injected in order to avoid overloading of the system. The area ratios in the isolated ion chromatograms for the parent ions in the QTOF were created as following: for the native peptide m/z 571.3 ($[M+2H]^{2+}$) and for the labelled peptide m/z 574.8 ($[M+2H]^{2+}$) was monitored (here the monoisotopic mass was considered). The area ratios were then compared to the data acquired on the LCQ of

the same theoretical ratios. Performing a two-sided paired *t*-test, there was no significant difference ($p>0.4$) between the data acquired on the different instruments.

Table 5.6: Comparison of area ratios (CYP1A2/internal standard peptide, ± standard error of the means) between QTOF and LCQ instruments. To in-solution digests of CYP1A2, different amounts of internal standard peptide were added and analysed using LC/MS/MS. There was no significant (paired *t*-test, $p>0.4$) difference between the instruments.

CYP1A2 in-solution digest injected [pmol]	**CYP1A2 specific labelled peptide injected [pmol]**	**Theoretical ratio (CYP1A2/internal standard peptide) [pmol/pmol]**	**Area ratio (CYP1A2/internal standard peptide)**	
			QTOF	**LCQ**
50	100	0.5	0.40 ± 0.03	0.38 ± 0.03
100	100	1	0.51 ± 0.02	0.55 ± 0.04
200	100	2	0.73 ± 0.01	0.78 ± 0.05

5.11 Quantification of Cytochrome P450 enzymes in liver microsomal samples

The MRM method to quantify CYP1A2, was applied to the quantification of this CYP protein in three biological samples. Liver microsomal preparations of three individuals were used here. The maximum of total protein amount acceptable was applied on gel and gel pieces were digested as described in 2.2.8.2 in the presence of 4 pmol internal standard for CYP1A2. Of each sample three parallel lanes were cut and each digest injected three times, so a maximum of nine injections were possible per individual liver sample. The CYP1A2 MRM-method was applied and between samples two blanks of isopropanol were injected to clean up the system. Table 5.7 shows the total protein amounts loaded onto the gel. Prior to the actual quantification experiment, samples were submitted to MS/MS analysis for identification and to confirm that the right band area was chosen for quantification. CYP1A2 was identified in all samples. For the CYP1A2 analysis it was important that the calculated amounts were in the range of quantities used for in-gel calibration. As pointed out in 5.7, the ratio-correlation followed a polynomial rather than a linear trend and it would be unreasonable to extrapolate how data would behave in a bigger range. Thus, results between 3 pmol and 7 pmol CYP1A2 on gel would correspond to the conclusions from the in-gel calibration

experiment 5.7. The CYP1A2 quantities shown in Table 5.7 corresponded to 3.4–6.9 pmol CYP1A2 enzyme on gel, which would be correspondent to the linear range determined in the in-gel digestion calibration experiment (5.7). More details about these liver samples can be found in appendix B. CYP1A2 in relation to microsomal protein was 165 pmol/mg for patient H6, 241 pmol/mg microsomal protein for patient H7 and in the case of patient H8 263 pmol/mg microsomal protein were found. The statistics for patient H7 and H8 were acceptable, but for individual H6, the CV value of 40% derived from bigger variation among the data. It would be possible that there is some variation in CYP content among individuals as investigated earlier [161]. Comparing the here-determined values with the ones in the study by Shimada *et al.* [217], there were an average of 42 pmol CYP1A2 per mg of microsomal protein detected immunochemically in human liver microsomes. Rodrigues *et al.* [216] determined similar amounts by immunoblotting (average of 45 pmol/mg total protein) whereas Guengerich and Turvy [259] found even less CYP1A2 by Western blot (average value of 26.5 pmol/mg microsomal protein). As observed with CYP2E1 (see section 4.7), the amounts found in this study were higher than those found elsewhere, although still within the same order of magnitude. The explanation for these differences for CYP1A2 are the same as those given above for CYP2E1, in particular inter-person variability and differences in sampling techniques, sample processing and methodology.

Table 5.7: Total protein content determined via the Bradford assay as well as protein amount used for SDS-PAGE of human liver microsomal samples of individual H6, H7 and H8. CYP1A2 quantification results are given in pmol CYP1A2 per mg total microsomal protein. N are the number of accepted sample injections.

Sample	Total protein content [µg/µl]	Amount loaded onto gel [µg]	CYP1A2/mg microsomal protein [pmol]	CV [%]	N
H6	5.45	24.53	165.43±66.75	40	9
H7	7.92	27.72	240.76±25.53	11	9
H8	4.91	24.55	262.6±76.18	29	9

5.12 Parallel quantification of proteins using MRM

The usage of internal standard peptides in quantitative proteomics has the ultimate goal of quantifying a multitude of proteins in one sample using one unique internal standard peptide per protein and its specific fragmentations. Therefore we had to test whether this was possible with our methods as well. A MRM method was set up to quantify CYP1A2 and CYP3A4. Thus, a peptide unique to CYP3A4 (EVTNFL*R) was used and the CYP1A2 quantification method was extended to two more scan events; scanning for the fragmentations of the light CYP3A4 peptide and the labelled internal standard CYP3A4 peptide respectively. This way it was possible to monitor the fragmentation of both proteins and thus quantifying two proteins in one run. However the validation with recombinant CYP3A4 protein did not lead to satisfying results and is not discussed here.

5.13 Resume and discussion

5.13.1 In-solution digest calibration and trial

For in-solution digest of proteins, standard protocols propose to use a trypsin/protein-ratio of 1/25–1/100 (weight/weight) [260]. There was no significant difference detected for different trypsin ratios used for an in-solution digest of CYP1A2. The data showed that with increasing amounts of injected equimolar protein digest/standard peptide-mixtures, the area of the internal standard increased much stronger than the area of the native peptide (Figure 5.9). This must be due to low solubilisation of recombinant CYP1A2 in the digestion, which means that even higher injection of this protein will not increase the signal to the same extent as the better solubilised internal standard peptide.

As experimental conditions such as pH and temperature were all the same and corresponding to standard recommendations, there might be a factor influencing the recovery of CYP1A2 in these experiments, which still has to be determined. As the results on the QTOF showed no significant difference to the ones acquired on the LCQ, this can also be seen as a proof in the assumption that the complications experienced are due to the protein itself and not the mass spectrometer used.

5.13.2 In-gel digest calibration and trial

The in-gel digestion of recombinant protein showed that a linear relationship was only seen for a small range of injected sample, i.e. corresponding to 3 to maximal 7 pmol CYP1A2 on the gel. The data for lower injection did not fulfil the required signal to noise ratio nor minimal signal intensity of 10,000 cps. For higher injections of recombinant CYP1A2 in-gel digest, the nano LC and ion-trap were overloaded and there was no longer a linear relationship verifiable. A polynomial function described the correlation of theoretical and experimental data better.

Recommended trypsin ratio for in-gel digestion of proteins is traditionally given as 1:25 [260] , which would mean in the case of 290 ng protein used, that 12 ng of trypsin would be needed. From this point of view, the deployed amounts here presented a clear excess of trypsin. However, for lower amounts of proteins it has been suggested to calculate the trypsin volume dependent on the gel-volume [260]. The digestion depends also on the density of the gel piece; higher protein density enhances digestion, lower density makes it difficult for the trypsin to reach all of the protein. Therefore 0.5 µg of trypsin per 15 mm^3 of gel piece are recommended [260]. The employed gel pieces with recombinant CYP1A2 bands had a thickness of 1 mm, were approximately 8 mm wide and between 2 and 3 mm broad. Following the above-mentioned recommendations, 500–750 ng would correspond to these numbers. Hence 1000 ng of trypsin was expected to be a sufficient amount. The comparison of the three trypsin amounts used in the in-gel-digestion trial, showed no significant difference.

5.14 Internal standard peptide choice and design

The quantification of CYP2E1 and the experience gathered here with CYP1A2 and CYP3A4 has shown, that there are several important factors to be considered when designing a method. Therefore synthetic peptides used as internal standards for peptide quantification have to be chosen carefully in order to guarantee successful analysis. The following workflow was used:

1. Selection of appropriate instrument
2. Analysis of tryptic digests of target protein using an exploratory MS/MS method (DDA) and identification of suitable peptides.
3. Selection of peptides unique for each protein.

4. Arrangement of peptides according to
    a) Amino acid sequence: chemically reactive residues (C, M, W) and chemically unstable sequences (N-G, N-terminal Q or N) should be eliminated.
    b) Hydrophobicity: hydrophobic peptides should be avoided
5. Choice of labelling position; label difference should be high enough to be able to differentiate peaks in mass spectrum and be preferable in a position, so that the fragment used for monitoring also contains the labelled amino acid.

First, experimental data from MS/MS experiments (using DDA) of human liver microsomes were analysed and screened for present CYPs. Only CYPs, which could be identified with the established MS/MS method, were selected. A list of detectable CYPs found in four liver microsomes (H1N to H4N) using the 1D-Gel–LC-MS/MS approach explained earlier (chapter 4) was compiled. In addition to this, liver control microsomes of three more individuals (H6,H7 and H8, see Table 5.7) were analysed using MS/MS as described in chapter 4.

In addition to the ability to detect and identify a protein and its peptides by a given instrument/method there are further requirements for an internal standard. Among CYPs there is a high sequence homology, especially among a family (>40%) or a subfamily (>55%). Therefore it is crucial to find unique peptides specific for the protein of interest. This way a distinction between enzyme-family members was possible. In order to decide about uniqueness of a peptide to a particular CYP, a Blast search (Basic Local Alignment Search Tool) was carried out for each peptide. Thus, using the blastp (protein-protein blast) algorithm, the identified internal standard peptide candidates were searched against a database of non-redundant (nr) human protein sequences. This way sequence similarities with other proteins could be discovered and peptides not showing identity over the entire sequence were eliminated from the list. As a result, 16 proteins and 81 peptides initially identified by LC-MS from liver microsomes were reduced to 10 proteins with 37 peptides unique to a particular CYP (Table 5.8). The list of unique peptides was then compared to data of human liver microsomes acquired on the same instrument previously ([200, 245]) and the peptide sequences mentioned there compared with the actual work. This step was added to ensure reproducibility, as the eventually designed peptide will be used in future studies. In addition to uniqueness a

suitable internal standard peptide has to show experiment-friendly physical-chemical properties to facilitate analysis. Various sequences or amino acids in a peptide can lead to complications; special attention has to be drawn to chemically reactive amino acids such as cysteine, methionine and tryptophan. They are susceptible to derivatisation and/or binding to other amino acids (sigma-genosys.eu.com/peptide_aqua.asp). Therefore it is suggested to select sequences without these amino acids ('good candidates', Table 5.8) Furthermore, the hydrophobicity of a peptide needs to be evaluated. The presence or absence of hydrophilic structures cannot only influence the solubility of a substance but also its behaviour when separated by LC or separated via gel-electrophoresis. It has to be therefore taken care when many hydrophobic amino acids such as alanine, phenylalanine, tryptophan or valine are incorporated in the desired sequence [152]. In practice it is very difficult to find sequences without hydrophobic amino acids.

It also has to be kept in mind that Leucine and Isoleucine in a peptide sequence may be able to differentiate between two isoforms of a protein, however due to their equal mass they cannot be differentiated.

After consideration of all requirements a list of six peptides as potential internal standard peptides for five CYPs were found. These peptides are the most suitable among the list identified in the samples. There are more peptide sequences possible, these 'optimised candidates' (Table 5.8) have been adapted according to experimental data.

Table 5.8: Summary of peptide candidates: 1Uniqueness was confirmed by performing a Blast Search (http://expasy.org/tools/blast/) against a human protein database. [2]Comparison with two other individual studies on human liver microsomes ([200, 245]). [3]Peptides containing chemically reactive residues (C, M, W) and chemically unstable sequences (N-G, N-term Q or N-term N) were eliminated. [4]Separation according to hydrophobicity results in theoretically 6 optimal candidates.

CYP450	Amino acid sequence observed	Position in protein	Unique peptides[1]	Unique &reproducible[2]	Good candidates[3]	Optimised candidates[4]
CYP2E1	GDLPAFHAHR	101-110	x	x	x	
	GIIFNNGPTWK	113-123	x			
	FKPEHFLNENGK	409-420	x	x		
	EALLDYKDEFSGR	88-100	x	x	x	
	FITLVPSNLPHEATR	360-374	x	x	x	x
	VKEHHQSLDPNCPR	250-263	x			
	DLTDCLLVEMEKEK	264-277	x			
	GTVVVPTLDSVLYDNQEFPDPEK	386-408	x	x	x	
	YGLLILMK	310-317	x	x		
	EAHFLLEALR	150-159	x	x	x	
	YSDYFKPFSTGK	423-434	x	x	x	
CYP2C9	LPPGPTPLPVIGNILQIGIK	29-48	x	x	x	x
	SHMPYTDAVVHEVQR	343-357	x	x		
CYP3A4	SAISIAEDEEWKR	116-128	x	x		
	LSLGGLLQPEKPVVLK	477-492	x	x	x	
	VWGFYDGQQPVLAITDPDMIK	71-91	x	x		
	APPTYDTVLQMEYLDMVVNETLR	343-365	x	x		
CYP4A11	VATALTLLR	469-477	x			
CYP3A5	LQKEIDAVLPNK	331-342	x	x		
	SAISLAEDEEWKR	116-141	x	x		
CYP2F1	EALVDQGEEFSGR	86-98	x			
CYP1A2	IGSTPVLVLSR	80-90	x	x	x	x
	DTTLNGFYIPK	393-403	x	x		
	FLTADGTAINKPLSEK	432-447	x	x		
	DITGALFK	282-289	x			
	TVQEHYQDFDKNSVR	267-277	x	x	x	x
	NTHEFVETASSGNPLDFFPILR	222-243	x	x	x	
	TVQEHYQDFDK	267-277	x	x	x	
CYP2C19	SNYFMPFSAGK	422-432	x			
CYP2C8	EALIDNGEEFSGR	85-97	x	x		
	DQNFLTLMK	191-199	x	x		
	VQEEIDHVIGR	323-333	x	x	x	x
	VQEEAHCLVEELR	145-157	x			
	VKEHQASLDVNNPR	248-261	x			
CYP2A6	GTEVYPMLGSVLR	388-400	x			
	GTGGANIDPTFFLSR	162-176	x	x	x	x
	DPSFFSNPQDFNPQHFLNEK	401-420	x	x	x	x

Chapter 6 Discussion and future work

In this work, our particular interest was the identification and quantification of metabolising enzymes in human microsomal samples. A major group of enzymes present in liver microsomes are the CYPs, which are the subject of the current investigations in terms of both identification and quantification. As a simple "yes"- or "no"-answer about their presence in a sample is not sufficient to provide information about their abundance and thereby the degree of their involvement in physiological processes, it is important to investigate their quantity in healthy as well as diseased states. This will not only provide further information about cellular metabolism but may also help to develop new therapeutic targets.

6.1 Protein identification

Protein identification was performed on human liver and lung microsomal samples using data dependent LC-MS/MS methods on two different nano-spray instruments, an ion trap and a Q-TOF mass analyser. The sample preparation prior to mass spectrometry analysis played an important role in the protein identification in these samples. Due to the fact that a tissue consists of a magnitude of different cell types, each consisting of various components with different protein expression patterns, unprocessed tissue samples present a very complex mixture containing an enormous number of proteins [261]. Therefore it was important to decrease sample complexity by e.g. looking at smaller substructures within cells [104]. Subcellular fractionation was applied to obtain the microsomal fraction, as the analysis of metabolising enzymes, especially CYPs was of particular interest. Protein separation was achieved by gel electrophoresis and at the peptide level by liquid chromatography. Nevertheless, due to differences in abundances of CYPs in different tissues, the limitations of the LC-MS/MS method [243], which was initially successful in the identification of CYPs in liver microsomes, were evident: submission of lung microsomal samples to a data dependent LC-MS/MS analysis failed to detect CYPs (exceptCYP5A1) despite various separation steps as mentioned above. It

was concluded that an additional "separation step", namely in the setup of the mass spectrometry method, was required to increase the number of identified proteins. To achieve this, the original data dependent LC-MS/MS method including a MS- survey scan in the mass range of 400–1500 *m/z* was subdivided by using three smaller mass ranges and each sample was analysed three times, once for each range (400–600, 600–800 and 800–1000 *m/z*). Thereby, 1.5 fold more proteins were detected. This showed that this was a suitable method to increase the sensitivity of the analysis. The CYP content of lung microsomes is known to be only about 10% of that of liver microsomes, but the improvement of the LC-MS/MS method did not help in identifying this special group of proteins in the lung samples tested. This showed that not only the total number of detectable proteins is a factor for evaluation of the outcome of a proteomic experiment, but also the dynamic range. As proteins are present in a wide dynamic range in biological samples [262] and high abundance proteins are detected predominantly by LC-MS/MS methods [79], the lower abundance proteins need more attention. It is known that the concentration of proteins in the well-characterised human plasma proteome vary by more then ten orders in magnitude [81]. A dynamic range of 10^{10} is a challenge for most methods and normally results in a bias towards more abundant proteins. The method used here relied on an automated data-dependent algorithm, which selected the most abundant signals for further analysis, resulting in a bias against lower abundance proteins. To reduce this bias, the original mass range investigated was split into three different ranges and each sample was analysed three times. Although this improved the number of proteins detected significantly (473 instead of 304), it still does not ensure a complete coverage, as for co-eluting peptides only the most abundant ones are being analysed. It also appeared that most of the detected peptides in the QTOF experiments were in the *m/z* range 400–800 and the method could be improved by investigating this range in greater depth. This shows also, that the digestion of proteins into peptides presents a complex question in itself. To address this, separation could be improved via two dimensional liquid chromatography, which has been performed successfully for mouse liver and lung CYPs [263]. As observed in our experiment, the use of several analytical runs with a reduced mass range increased the number of identified proteins, but also some proteins were not identified in the reduced mass range runs, but were seen with the original large mass range method. Liu *et al.* [264] have investigated how

random is the detection of a peptide in a data dependent MS/MS analysis. The origin for their research was the observation that the replicate analysis of a sample, using the same method, resulted in more identifications, as in each run some new proteins were found. In their study they found for example that for an almost complete (95%) identification of all proteins in a yeast cell lysate, 10 analyses would be necessary.

A further possibility to increase protein identification would be to identify the most intense signals in the mass spectrum and then exclude these ones from analysis in a second MS/MS run, so lower abundance peptides would be detected; this could then be repeated several times. This tactic has been suggested in a study by Chen *et al.* [265], where in a mode of extended dynamic exclusion, a list of peptide precursor ions for rejection was created for an experiment on LC-MALDI. After initial LC-MS/MS analysis and database search a list with the most positively identified peptides was made. For consecutive runs this list was used as an exclusion list, and peptides positively identified in the next run were added to the list for the following run. This way, Chen *at al.* managed to improve the identification in an *E. coli* lysate by 25% compared to double analysis of the same sample. Also the preparation of gel samples could be changed to cut more bands as illustrated in the work of Petushkova *et al.* [255], who merge the data of 20 or more pieces together when carrying out database search. However, all these methods require much longer experimental steps and analysis times on the instrument; for a method to be used for routine analysis, a compromise between data acquisition time and amount of detected data has to be made. The alternatives would be to use different instruments, with better sensitivity, higher data throughput, different techniques, or additional sample preparation. A prominent example is the proteomic analysis of blood, which is complicated by blood-typical components, such as albumin and immunoglobulins and other high abundant proteins [266]. Depletion of these high abundance proteins from plasma samples is known to help to identify lower abundance proteins which would be demanding to detect otherwise [267]. However, Granger *et al.* [268] found that with depletion of plasma, low abundant proteins such as cytokines, were also removed, presumably due to unspecific binding to the column matrix and albumin; the losses caused by depletion of a sample should not be underestimated [268]. As described over twenty years ago by Guengerich *et al.* [269] , additional solvents and detergents could be used in order to isolate specific protein families, e.g. CYPs from microsomes, so an

additional separation step could be employed. These detergents in turn could have an influence on the performance of further analysis such as gel electrophoresis, digestion as well as detection in mass spectrometry and continuous research is being undertaken in this area [196, 262, 270, 271]. In the case of microsomal CYPs, which are equipped with a membrane anchor [196], the same obstacles commonly observed with membrane proteins are applicable [214, 235]. For example, the 2D-gel analysis of microsomes was unsatisfactory in regard of detection of CYPs [214, 255]. In addition to the issues regarding detection of membrane proteins via two dimensional gel electrophoresis, it was also shown earlier that this technique cannot guarantee the detection of low abundance proteins [262]. Thus, as mentioned above, the challenge in proteomics will be to overcome the limitations which are still there due to the large dynamic range in the human proteome in comparison to the relatively lower dynamic range of MS analysis methods, which are quoted to be around 10^4 to 10^6 at best when combined with on-line separations techniques [270].

It should not be forgotten that during sample preparation there could also be losses, which have to be taken into account as well. From the first moment of handling a sample, caution has to be taken, as for example metabolite levels in human tissue were found to change in minutes or seconds after dissection from the body [261]. Studies on the comparison of different tissue sources for microsomal preparation have been undertaken by Song *et al.* [271]. When they compared fresh liver, homogenized frozen liver and frozen liver, they found that fresh liver or liver homogenised and then frozen were the best starting material, whereas frozen liver tissue was not recommended to be used for microsomal preparation. This study was performed with mouse liver samples which were freely available and could give different results for human samples which are more difficult to collect; however, it showed that every detail could have an influence on the outcome of a proteomic analysis. It has also been found, that CYPs were prevalent in the perivenous region of the mammalian liver acinus [272], which implies the importance of knowledge about the part of a particular tissue from which a sample is taken. Therefore it was proposed to develop standard procedures for sample handling and storage [261]. This would be important in terms of comparability of results obtained in different laboratories, but it poses obstacles where human tissue is provided from

surgery, where the main priority is patient health rather than a standard experimental protocol.

Using the data dependent method with a survey scan over a large mass range, no CYPs could be detected in liver tumour samples. We cannot conclude from our results that no CYPs are present in the tumour samples, but one would rather assume that they are below the limit of detection. For these samples, we additionally used a MRM method, looking for specific CYPs. A method developed for the selective monitoring of a CYP2E1 specific tryptic peptide supported by a synthetic stable isotope labelled standard peptide specific to CYP2E1 was employed and no signal for the endogenous peptide was detected in the sample. The limit of quantification was determined to be 1 pmol CYP2E1 on gel, so we can assume with these findings that the level of CYP2E1 in these tumour microsomes was below 1 pmol on gel, which corresponds to 40 pmol/mg microsomal protein, based on 25 µg average total micromsomal protein application on gel.

A major obstacle in the determination of CYP proteins is the fact that they are membrane bound and the proteomic analysis of membrane or membrane associated proteins remains a challenge. However, of the proteins identified in our study of lung microsomal samples, 62 % were annotated to the membrane fractions, such as NADPH-cytochrome P450 reductase, dolichly-diphosphooligosaccharide-protein glycosyltransferase, aldehyde dehydrogenase, glutathione S-transferase cytochrome *c* oxidase subunit 2, COX2, epoxide hydrolase, Na/K-ATPase alpha 1 subunit, estradiol 17β-dehydrogenase and amine oxidase B. This shows that the method was capable of detecting membrane proteins. The failure to detect CYPs is therefore likely a consequence of their low abundance in lung microsomes. According to Shimada *et al.* [235] the spectroscopic quantification of the total CYP amount in lung microsomes provided a concentration of about 10 pmol/mg microsomal protein. However, taking into account our quantification results in the liver, where we determined CYP2E1 in four individuals to range between 88 and 200 pmol/mg microsomal protein, 10 pmol/mg total CYP protein is clearly below the limit of detection for this method, i.e. 40 pmol/mg microsomal protein. For the routine identification of lung CYPs, it is therefore necessary to improve the existing method to increase its sensitivity by e.g. further enrichment. Apart from introducing standards for sample acquisition and preparation, more

sophisticated data-dependent analysis algorithms for analysing the data from survey scans to exclude abundant proteins in second and third analysis runs – in specific chromatographic time windows – would help to achieve a higher sensitivity without modification of the original method.

6.2 Protein quantification

In this work, human liver microsomal CYPs were quantified using a stable isotope labelled internal standard-quantification method. As the identification with mass spectrometry had been successful, the application of this technique for the purpose of quantification was appropriate [214]. Various approaches have been undertaken previously to measure the CYP content in liver microsomes on the protein level using immunochemical as well as mass spectrometry/proteomic methods. The complication of using measurements on the level of mRNA was described above (1.4). It is difficult to compare the results acquired with the different methods, as each has got its advantages and disadvantages. In terms of proteomic quantitative experiments, there is still little data available and most studies on the CYP-protein level so far deal with immunochemical detection [216, 217]. This always implies the availability of a monoclonal antibody for each CYP enzyme. With the high sequence homology among CYPs it can be difficult to find and produce monoclonal antibodies for each CYP [273]. At the moment, quantitative proteomics techniques incorporating stable isotopes are commonly used either for relative or absolute quantification [151]. Stable isotopes have been employed in a variety of quantitative proteomics studies for CYP analysis [257]. The application of stable isotope labelled internal standard peptides for quantification of CYPs was chosen here as this offered a way towards absolute quantification. Other isotope labelling methods used by others have been directed at the the comparison of two sets of samples labelled differently in order to receive relatative, rather than absolute, quantitative information. As the sequence homology between CYPs is very high, it is necessary to look for unique peptides representing a peptide sequence, which can only be found in one particular CYP and is preferably accessible via enzymatic digest, such as trypsin. The advantage of a unique isotope labelled peptide is that it can mimic the properties of the endogenous analyte, as the only difference between native and synthetic peptide is expected to be the molecular weight. The idea of unique

peptides for each CYP was also addressed by Kornilayev and Altermann [274], who developed a combination of mass spectrometry and immunochemistry techniques to differentiate between CYP forms. The combination is enabled via polyclonal antibodies specific for unique peptide sequences of individual CYPs, which absence or presence can be detected via mass spectrometry as well as ELISA. It can be seen as an advantage that the results derive from two different experiments and can confirm one another. This might in future be extended for quantification of CYPs as well. In another study Alterman *et al.* [275] used unique synthetic peptides without any labelling as internal standards for human and rat liver CYPs. However, this methodology is suspect as the standard is chemically different from the analyte.

The above discussion illustrates that in addition to our work, others have taken the approach to find unique targets for each CYP, which can then be used in quantitative experiments. This approach contributes to the selectivity of quantitative experiments and also can be used in qualitative studies. Alterman *et al.* [275] presented a list of possible peptides for three CYPs (1A2, 2E1 and 2C19), which they chose according to their criteria of applicability to function as an internal standard for mass spectrometry analysis. We also created a list of ideal candidates for several CYPs (Table 5.8) according to our results and experiences. This is the first step into creating a global list of unique peptide candidates for various CYP proteins [276].

Although the results obtained in this study for CYP2E1 and CYP1A2 quantification were higher than those reported elsewhere, they were still in the same order of magnitude and differences can be attributed to inter-individual differences in protein expression (see 6.1), differences in methodology, sample preparation and sample processing.

Table 6.1 presents the results of four different studies, which quantified CYPs either immunochemically or using stable isotopes. The results in our work were derived from a stable isotope internal standard experiment, which were not employed for the analysis of human liver microsomes previously. Therefore, the comparisons with other studies means the comparison of different methodologies as well as sample populations. To our knowledge only three previous immunochemical studies were published on human liver microsomes, which all detected lower amounts of CYP1A2 and CYP2E1 than in our study. The ICAT method was carried out on rat microsomes, so they cannot

be compared directly to our results and are only stated for reference. As mentioned above, large inter-individual differences in CYP expression can occur [161]. The tendency to detect higher enzyme levels in our study using the stable isotope internal standard method could be due to its increased selectivity and sensitivity, achieved through the utilisation of multiple reaction monitoring. The probability to detect a different protein was smaller than with immunochemical methods, especially with the high degree of sequence homology among members of CYP subfamilies.

Table 6.1: Overview of liver microsomal quantification of CYP1A2 and CYP2E1 of immunochemical and stable isotope methods in literature and work presented here. CYP contents are stated as pmol/mg microsomal protein.

Study by	**Type of experiment**	**Liver microsomal organism**	**CYP1A2**	**CYP2E1**
This work	Stable isotope dilution	human	165–263	88–200
Shimada *et al.* [217]	Immunochemical	human	42	22
Rodrigues *et al.* [216]	Immunochemical	human	45	49
Guengerich and Turvy[259]	Immunochemical	human	119	26.5
Jenkins [257]	ICAT	rat	35.13	5.38

Some of the difficulties experienced in our work might have been due to the fact that CYPs are hydrophobic membrane associated proteins, which are quantified after an in-gel digest was performed. In the case of the CYP2E1 peptide, the hydrophobicity of the internal standard peptide became an issue, as the solvents and the LC-gradients used had to be adapted. Also the low recoveries of CYP1A2 in in-solution digests was probably due to problems with protein solubilistaion.

The deployment of an internal standard, added together with trypsin, is still optimal to overcome losses possibly caused during sample handling and after digestion. Losses are possible during destaining of bands, during digestion and extraction, mainly caused by sticking to surfaces in sample tubes or peptide tips [277], even though specially treated equipment was used (such as siliconised sample tubes). The in-gel digestion coupled to LC-MS/MS analysis has proved successfully to identify CYPs in human liver [243] and as explained above, it is necessary to separate the complex microsomal samples, so this was done here via gel-electrophoresis. As well the extraction efficiency, the

digestion of the protein is an important factor for recovery and has therefore to be complete. This cannot be compensated with an internal standard peptide, which is already present in a tryptic peptide. We investigated digestion efficiency of CYP1A2 and tested in-solution as well as in-gel digestion efficiency. For in-solution digest experiments, no significant influence on the amounts of trypsin added could be determined. However, it was observed, that injections of increasing amounts of equimolar protein-digest/isotope-labelled peptide mixtures lead to decreasing ratios, as the area of the internal standard peptide increased more strongly than the area of the native peptide and thereby implicated low solubility of the recombinant protein in this digestion experiments. The reason for this could not be explained, neither by trypsin trials, investigating digestion time and enzyme amounts added nor by trying to increase solubility of the proteins; it seemed to be rather related to the protein properties itself than the method. Future studies could be carried out by adding commercial surfactants designed to assist proteolysis. Similarly, trials for the in-gel digest showed no significant differences for the usage of different trypsin amounts.

The isotope dilution method is designed to compensate for sample losses during sample preparation. However, sample losses are possible during the entire protocol, and can be due for example to adsorption, degradation or interaction with other proteins. A consequence of losses on the protein level would be a low recovery of recombiant protein. Recovery rates are an issue which need to be considered and it has to be explained how sample losses occurred. The nature of the isotope labelled peptide does not allow its addition at a step preceding digestion, as the microsomal sample is first submitted to gel-electrophoresis and the peptide and protein would not coelute in the same band on the gel due to the difference in their molecular weight. Also, the isotope labelled peptide is a surrogate for the trypsin digested protein not the intact protein. In order to better simulate digestion and compensate for incompletion, Barnidge *et al.* [278] tested the usage of extended internal standard peptides which had to be digested also to obtain the actual internal standard. Another method was introduced by Beynon *et al.* [279]; they expressed the gene sequence of several internal standard peptides in a bacterial culture; after expression and purification, the expressed "peptide-chain" was digested together with the analyte and the individual peptide sequences were released and analysis was performed. The expression of the recombinant protein was performed on

stable isotope enriched medium, so the peptides contained automatically the label. Thereby several proteins could be quantified simultaneously and digestion efficiency was assumed to be the same for analyte and standard. However, a disadvantage of this method is, that in theory the quantity of each standard has to be determined after each expression using amino acid analysis, in order to guarantee correct results. Similarly, the assumption that two entirely different proteins give the same digestion efficiency is not necessary valid. Despite economical and technical effort to synthesise whole proteins, there have been attempts to employ whole protein-internal standards. The crucial factor was the observations mentioned above that digestion efficiency and recovery are pivotal issues to be considered. Ji *et al.* [280] report the quantification of a plasma protein using a ^{15}N-labelled version of the analyte. Accuracy and accordingly recovery values varied between 101.5% and 104.7%. Another tactic is the modification of a protein, for example through carbamidomethylation [281], biotynilation [282] or N-terminal truncation [283]; with the resultant mass differences creating internal standards of the analytes. Duan *et al.* [284]used *p*-CPB sulfuhydryl alkylation of cysteines for CYP protein quantification in rat liver microsomes. They determined e. g. CYP2E1 as 134 pmol/mg microsomal protein. This was almost four times higher compared to the results of Jenkins *et al.* [257], who stated a CYP2E1 content of 35.13 pmol/mg of rat liver microsomal protein. However, it is still important to monitor the synthesis of these modified proteins in order to know their exact mass. The chemical modification also implies a change in the physicochemical behaviour and can thereby seen as a draw–back of this technique. For recombinant protein-standards, it is also always necessary to perform an accurate amino acid analysis in order to know the quantity of standard added. The complete incorporation of the stable isotopes (in stable isotope enriched growth medium) also has to be proved. Label free approaches [285] have been reported as well, mostly carried out on high-resolution instruments, such as the Orbitrap or LTQ-FT mass analysers [286]. Nevertheless, these different approaches would still be a challenge if applied to CYPs, as their sequence homology would make it difficult to find suitable standards, in order to quantify CYP isoforms selectively out of complex mixtures. Another promising approach is the use of surrogate standards or "pseudo internal standards", as described by Tabata *et al.* [287], where housekeeping proteins, i. e. proteins known not to change

expression levels, are used as internal standard to evaluate abundances via calculation of ratios between surrogate standards and analyte.

A further issue, which has to be addressed, was the observed non-linearity of calibration curves especially in the higher concentrations range, especially for the in-gel digestion of CYP1A2 (chapter5). In the area of small molecules quantification via isotope dilution, this subject has been discussed extensively [288] and possible explanations were found in detector saturation if high amounts are injected [289], as well as potential dimer or multimer formation in the gas-phase [290]. Non linearity could also be related to droplet saturation in electrospray droplet formation at high concentrations [50]. These findings must not be neglected and on the contrary it has to be investigated closly whether these phenomenons also apply to protein-mass spectrometry. Damen *et al.* [291], when quantifying a whole protein, made similar observations and linearity problems could be resolved there by using another solvent and changing the source temperature. Looking at proteins and peptides as large molecules, consisting of amino acids, which can possess various chemical moieties, it could be possible that there are diverse interactions going on in the formation of charged ions in the mass spectrometry interface. Little is published yet about these difficulties in proteomics, but with more attempts to quantify proteins and the aim to deliver correct results, this will be inevitable.

6.3 Future work

We have shown that mass spectrometry is a powerful tool for the identification and quantification of CYPs. The analysis of this super family will remain an important issue, in particular because of their rôle in metabolising drugs and other xenobiotic compounds. It is necessary to give consideration to common challenges in proteomic research such as high dynamic range of protein expression in biological samples and the comparatively smaller range of detection by mass spectrometry instruments. The introduction of unique peptides as internal standards in this work, presents a promising new technique, which not only improves sensitivity, but also increases selectivity. This is important for proteins like CYPs as a high sequence homology makes it difficult to differentiate between different members of subfamilies. The collection of unique peptides as candidates for possible internal standards as a database, where all CYPs –

and possibly other proteins – could be listed, would facilitate further research. This could be done for different organisms as well as different tissues. Regarding different tissues however, our experience with detection of CYPs in lung tissue has shown, that more work is required in order to enhance sensitivity and the limit of detection. In turn, the combination of multiple reaction monitoring (MRM) methods in identification experiments seems a suitable starting point for improvement. A further experiment that should also be performed is the investigation of the number of peptides or proteins respectively that can be quantified via multiple reaction monitoring in one single analytical run, in particular with regard to new mass analysers which are capable of monitoring large numbers of transitions in a short time. This could also mean that peptides have to be chosen according to their retention time. Thus, during liquid chromatographic separation, there would be many time frames of elution of different peptides and MS/MS fragmentation could be divided into these time frames as well. This would require more validation of the LC-performance in order to create a reliable method.

The sample types used for analysis may also need further thought, i.e. alternatives to tissue biopsies should be considered. In terms of principle studies, tissue samples, such as those derived from biopsies or during surgeries are useful to investigate the proteome of that particular tissue and e. g. explain the origin of a disease or malfunction. However, in fields such as biomarker monitoring and development, the analysis of blood or urine would be easier, as it is less invasive, quicker and more useful when supervising a course of disease or its treatment. Again, here the development of robust, high throughput routine-methods is of need. Proteomics science is a fast emerging field and major advances have been made over the last ten years, in particular regarding new instrumentations such as Orbitrap and FTICR that have achieved high sensitivity, accuracy and resolution and this will facilitate the development of suitable quantitative methods.

The subcellular fractionation as such also needs some further investigation. As it is necessary to isolate the microsomal fraction for CYP analysis, it would be important to know whether all CYPs naturally apparent in the microsomal membrane fraction of the cell were all extracted via the subcellular fractionation method. The here-created quantitative mass spectrometry method now would offer the possibility to give

quantitative evidence. The homogenisation and ultracentrifugation of liver tissue into its subcellular fractions had to be successful in order to avoid cross-contamination between fractions. Thus it is possible to quantify and state the content of all CYPs confidently in liver control samples. In theory all processed fractions have to be analysed. A first attempt was made here investigating potential loss between mitochondrial and microsomal fraction. In order to control this, also the mitochondrial fraction, which is separated in the step before the separation of the microsomal fraction, was examined and analysed for CYPs. In three investigated individuals, no CYPs were identified in the mitochondria fraction using a data dependent LC-MS/MS method.

The LC-MS/MS method used in this study can be used to compare expression patterns of metabolising enzymes in different tissues, tissue states (i. e. healthy, primary and secondary tumour) and following intervention with drugs. Combined with other 'omics' techniques, such as metabolomics and transcriptomics, this will allow a further understanding of cellular processes and the development of novel chemotherapeutic agents targeted directly to diseased tissue.

Bibliography

1. Huber, L. A. *Is Proteomics Heading in the Wrong Direction?* Nat Rev Mol Cell Biol **4**, 74-80 (2003).
2. Cravatt, B. F., Simon, G. M. and Yates Iii, J. R. *The Biological Impact of Mass-Spectrometry-Based Proteomics.* Nature **450**, 991-1000 (2007).
3. Tyers, M. and Mann, M. *From Genomics to Proteomics.* Nature **422**, 193-197 (2003).
4. Castell, J. V., Teresa Donato, M. and Gomez-Lechon, M. J. *Metabolism and Bioactivation of Toxicants in the Lung. The in Vitro Cellular Approach.* Exp Toxicol Pathol **57**, 189-204 (2005).
5. Aebersold, R. and Goodlett, D. R. *Mass Spectrometry in Proteomics.* Chem Rev **101**, 269-296 (2001).
6. Patterson, S. D. *Proteomics: Evolution of the Technology.* Biotechniques **35**, 440-444 (2003).
7. Steel, L. F., Haab, B. B. and Hanash, S. M. *Methods of Comparative Proteomic Profiling for Disease Diagnostics.* J Chromatogr B **815**, 275-284 (2005).
8. Lindsay, M. A. *Target Discovery.* Nat Rev Drug Discov **2**, 831-838 (2003).
9. Cornett, D. S., Reyzer, M. L., Chaurand, P. and Caprioli, R. M. *Maldi Imaging Mass Spectrometry: Molecular Snapshots of Biochemical Systems.* Nat Meth **4**, 828-833 (2007).
10. Cox, J. and Mann, M. *Is Proteomics the New Genomics?* Cell **130**, 395-398 (2007).
11. Hanash, S. *Hupo Initiatives Relevant to Clinical Proteomics.* Mol Cell Proteomics **3**, 298-301 (2004).
12. Hamacher, M., Apweiler, R., Arnold, G., Becker, A., Bluggel, M., Carrette, O., Colvis, C., Dunn, M. J., Frohlich, T., Fountoulakis, M., van Hall, A., Herberg, F., Ji, J., Kretzschmar, H., Lewczuk, P., Lubec, G., Marcus, K., Martens, L., Palacios Bustamante, N., Park, Y. M., Pennington, S. R., Robben, J., Stuhler, K., Reidegeld, K. A., Riederer, P., Rossier, J., Sanchez, J. C., Schrader, M., Stephan, C., Tagle, D., Thiele, H., Wang, J., Wiltfang, J., Yoo, J. S., Zhang, C., Klose, J. and Meyer, H. E. *Hupo Brain Proteome Project: Summary of the Pilot Phase and Introduction of a Comprehensive Data Reprocessing Strategy.* Proteomics **6**, 4890-4898 (2006).
13. Zheng, J., Gao, X., Beretta, L. and He, F. *The Human Liver Proteome Project (Hlpp) Workshop During the 4th Hupo World Congress.* Proteomics **6**, 1716-1718 (2006).
14. Omenn, G. S., States, D. J., Adamski, M., Blackwell, T. W., Menon, R., Hermjakob, H., Apweiler, R., Haab, B. B., Simpson, R. J., Eddes, J. S., Kapp, E. A., Moritz, R. L., Chan, D. W., Rai, A. J., Admon, A., Aebersold, R., Eng, J., Hancock, W. S., Hefta, S. A., Meyer, H., Paik, Y. K., Yoo, J. S., Ping, P., Pounds, J., Adkins, J., Qian, X., Wang, R., Wasinger, V., Wu, C. Y., Zhao, X., Zeng, R., Archakov, A., Tsugita, A., Beer, I., Pandey, A., Pisano, M., Andrews, P., Tammen, H., Speicher, D. W. and Hanash, S. M. *Overview of the Hupo Plasma*

Proteome Project: Results from the Pilot Phase with 35 collaborating Laboratories and Multiple Analytical Groups, Generating a Core Dataset of 3020 proteins and a Publicly-Available Database. Proteomics **5**, 3226-3245 (2005).

15. DeRisi, J. L., Iyer, V. R. and Brown, P. O. *Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale.* Science **278**, 680-686 (1997).
16. Cox, B., Kislinger, T. and Emili, A. *Integrating Gene and Protein Expression Data: Pattern Analysis and Profile Mining.* Methods **35**, 303-314 (2005).
17. Gygi, S. P., Rochon, Y., Franza, B. R. and Aebersold, R. *Correlation between Protein and Mrna Abundance in Yeast.* Mol Cell Biol **19**, 1720-1730 (1999).
18. Aebersold, R. and Mann, M. *Mass Spectrometry-Based Proteomics.* Nature **422**, 198-207 (2003).
19. Witze, E. S., Old, W. M., Resing, K. A. and Ahn, N. G. *Mapping Protein Post-Translational Modifications with Mass Spectrometry.* Nat Meth **4**, 798-806 (2007).
20. Shapiro, A. L., Vinuela, E. and V. Maizel, J. *Molecular Weight Estimation of Polypeptide Chains by Electrophoresis in Sds-Polyacrylamide Gels.* Biochem Biophys Res Commun **28**, 815-820 (1967).
21. Weber, K. and Osborn, M. *The Reliability of Molecular Weight Determinations by Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis.* J Biol Chem **244**, 4406-4412 (1969).
22. Granvogl, B., Plöscher, M. and Eichacker, L. *Sample Preparation by in-Gel Digestion for Mass Spectrometry-Based Proteomics.* Anal Bioanal Chem **389**, 991-1002 (2007).
23. Righetti, P. G. and Gelfi, C. *Electrophoresis Gel Media: The State of the Art.* J Chromatogr B Analyt Technol Biomed Life Sci **699**, 63-75 (1997).
24. Hames, B. D. *Gel Electrophoresis of Proteins: A Practical Approach* (IRL Press at Oxford University Press, 1990).
25. Ornstein, L. *Disc-Electrophoresis-I Background and Theory.* Ann. N. Y. Acad. Sci. **121**, 321-349 (1964).
26. O'Farrell, P. H. *High Resolution Two-Dimensional Electrophoresis of Proteins.* J Biol Chem **250**, 4007-4021 (1975).
27. Dunn, M. J. *Two-Dimensional Gel Electrophoresis of Proteins.* J Chromatogr B **418**, 145-185 (1987).
28. Santoni, V., Molloy, M. and Rabilloud, T. *Membrane Proteins and Proteomics: Un Amour Impossible?* Electrophoresis **21**, 1054-1070 (2000).
29. Galeva, N. and Altermann, M. *Comparison of One-Dimensional and Two-Dimensional Gel Electrophoresis as a Separation Tool for Proteomic Analysis of Rat Liver Microsomes: Cytochromes P450 and Other Membrane Proteins.* Proteomics **2**, 713-722 (2002).
30. Goldstein, E. Sitzungsberichte der königlich-preußischen Akademie der Wissenschaften **39**, 691-699 (1886).
31. Thomson, J. J. Philosophical Magazine **13**, 561 (1907).
32. Aston, F. W. *Neon.* Nature **104**, 334 (1919).
33. Aston, F. W. *A Positive Ray Spectrograph.* Philosophical Magazine **38**, 707 (1919).
34. Biemann, K. *Laying the Groundwork for Proteomics: Mass Spectrometry from 1958 to 1988.* Int J Mass Spectrom **259**, 1-7 (2007).
35. Barber, M., Bordoli, R. S., Garner, G. V., Gordon, D. B., Sedgwick, R. D., Tetler, L. W. and Tyler, A. N. *Fast-Atom-Bombardment Mass Spectra of Enkephalins.* Biochem J **197**, 401-404 (1981).

36. Karas, M. and Hillenkamp, F. *Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10,000 Daltons.* Anal Chem **60**, 2299-2301 (1988).
37. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M. *Electrospray Ionization for Mass Spectrometry of Large Biomolecules.* Science **246**, 64-71 (1989).
38. de Hoffmann, E. and Vincent, S. *Mass Spectrometry: Principles and Applications* (eds. de Hoffmann, E. and Vincent, S.) (John Wiley and Sons, Ltd, 2002).
39. Wells, J. M. and McLuckey, S. A. *Collision-Induced Dissociation (Cid) of Peptides and Proteins.* Methods Enzymol **402**, 148-185 (2005).
40. Meng, C. K., Mann, M. and Fenn, J. B. in 36th American Society for Mass Spectrometry Conference on Mass Spectrometry and Allied Topics 771-772 (ASMS, San Francisco,CA, 1988).
41. Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D. and Alice, M. B. *Molecular Beams of Macroions.* J Chem Phys **49**, 2240-2249 (1968).
42. Griffiths, W. J., Jonsson, A. P., Liu, S., Rai, D. K. and Wang, Y. *Electrospray and Tandem Mass Spectrometry in Biochemistry.* Biochem J **355**, 545-561 (2001).
43. Simon, J. G. *Electrospray: Principles and Practice.* J Mass Spectrom **32**, 677-688 (1997).
44. Taflin, D. C., Ward, T. L. and Davis, E. J. *Electrified Droplet Fission and the Rayleigh Limit.* Langmuir **5**, 376-384 (1989).
45. Mabbett, S. R., Zilch, L. W., Maze, J. T., Smith, J. W. and Jarrold, M. F. *Pulsed Acceleration Charge Detection Mass Spectrometry: Application to Weighing Electrosprayed Droplets.* Anal Chem **79**, 8431-8439 (2007).
46. Mack, L. L., Kralik, P., Rheude, A. and Dole, M. *Molecular Beams of Macroions. Ii.* J Chem Phys **52**, 4977-4986 (1970).
47. Iribarne, J. V. and Thomson, B. A. *On the Evaporation of Small Ions from Charged Droplets.* J Chem Phys **64**, 2287-2294 (1976).
48. Wilm, M. S. and Mann, M. *Electrospray and Taylor-Cone Theory, Dole's Beam of Macromolecules at Last?* Int J Mass Spectrom Ion Process **136**, 167-180 (1994).
49. Wilm, M. and Mann, M. *Analytical Properties of the Nanoelectrospray Ion Source.* Anal Chem **68**, 1-8 (1996).
50. Cech, N. B. and Enke, C. G. *Practical Implications of Some Recent Studies in Electrospray Ionization Fundamentals.* Mass Spectrom Rev **20**, 362-387 (2001).
51. Paul, W., Reinhard, H. P. and von Zahn, U. *Das Elektrische Massenfilter Als Massenspektrometer Und Isotopentrenner.* Z Phys A At Nucl **152**, 143-182 (1958).
52. Paul, W. *Electromagnetic Traps for Charged and Neutral Particles (Nobel Lecture).* Angew Chem Int Ed Engl **29**, 739-748 (1990).
53. Stafford, G. C., Kelley, P. E., Syka, J. E. P., Reynolds, W. E. and Todd, J. F. J. *Recent Improvements in and Analytical Applications of Advanced Ion Trap Technology.* Int J Mass Spectrom Ion Process **60**, 85-98 (1984).
54. Jonscher, K. R. and Yates, J. R. *The Quadrupole Ion Trap Mass Spectrometer--a Small Solution to a Big Challenge.* Anal Biochem **244**, 1-15 (1997).
55. Hager, J. W. *A New Linear Ion Trap Mass Spectrometer.* Rapid Commun Mass Spectrom **16**, 512-526 (2002).
56. John, R. Y., III. *Mass Spectrometry and the Age of the Proteome.* J Mass Spectrom **33**, 1-19 (1998).

57. Chernushevich, I. V., Loboda, A. V. and Thomson, B. A. T. *An Introduction to Quadrupole-Time-of-Flight Mass Spectrometry.* J Mass Spectrom **36**, 849-865 (2001).
58. Mamyrin, B. A., Karataev, V. I., Shmikk, D. V. and Zagulin, V. A. *The Mass-Reflectron, a New Nonmagnetic Time-of-Flight Mass Spectrometer with High Resolution.* . Zh Eksp Teor Fiz **64**, 82-89 (1973).
59. Johnson, R. S., Martin, S. A., Biemann, K., Stults, J. T. and Watson, J. T. *Novel Fragmentation Process of Peptides by Collision-Induced Decomposition in a Tandem Mass Spectrometer: Differentiation of Leucine and Isoleucine.* Anal Chem **59**, 2621-2625 (1987).
60. Wysocki, V. H., Tsaprailis, G., Smith, L. L. and Breci, L. A. *Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation.* J Mass Spectrom **35**, 1399-1406 (2000).
61. Dongre, A. R., Jones, J. L., Somogyi, A. and Wysocki, V. H. *Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model.* J Am Chem Soc **118**, 8365-8374 (1996).
62. Noga, M., Sucharski, F., Suder, P. and Silberring, J. *A Practical Guide to Nano-Lc Troubleshooting.* J Sep Sci **30**, 2179-2189 (2007).
63. Nesvizhskii, A. I. and Aebersold, R. *Analysis, Statistical Validation and Dissemination of Large-Scale Proteomics Datasets Generated by Tandem Ms.* Drug Discovery Today **9**, 173-181 (2004).
64. Perkins, D. N., Pappin, D. J. C., Creasy, D. M. and Cottrell, J. S. *Probability-Based Protein Identification by Searching Sequence Databases Using Mass Spectrometry Data.* Electrophoresis **20**, 3551-3567 (1999).
65. Eng, J. K., McCormack, A. L. and Yates, J. R. *An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database.* J Am Soc Mass Spectrom **5**, 976-989 (1994).
66. Clauser, K. R., Baker, P. and Burlingame, A. L. *Role of Accurate Mass Measurement (+/- 10 Ppm) in Protein Identification Strategies Employing Ms or Ms/Ms and Database Searching.* Anal Chem **71**, 2871-2882 (1999).
67. Pappin, D. J. C., Hojrup, P. and Bleasby, A. J. *Rapid Identification of Proteins by Peptide-Mass Fingerprinting.* Curr Biol **3**, 327-332 (1993).
68. Elias, J. E., Haas, W., Faherty, B. K. and Gygi, S. P. *Comparative Evaluation of Mass Spectrometry Platforms Used in Large-Scale Proteomics Investigations.* Nat Methods **2**, 667-675 (2005).
69. Carr, S., Aebersold, R., Baldwin, M., Burlingame, A., Clauser, K. and Nesvizhskii, A. *The Need for Guidelines in Publication of Peptide and Protein Identification Data: Working Group on Publication Guidelines for Peptide and Protein Identification Data.* Mol Cell Proteomics **3**, 531-533 (2004).
70. Beer, I., Barnea, E., Ziv, T. and Admon, A. *Improving Large-Scale Proteomics by Clustering of Mass Spectrometry Data.* Proteomics **4**, 950-960 (2004).
71. Moore, R. E., Young, M. K. and Lee, T. D. *Method for Screening Peptide Fragment Ion Mass Spectra Prior to Database Searching.* J Am Soc Mass Spectrom **11**, 422-426 (2000).
72. Yang, X., Dondeti, V., Dezube, R., Maynard, D. M., Geer, L. Y., Epstein, J., Chen, X., Markey, S. P. and Kowalak, J. A. *Dbparser: Web-Based Software for Shotgun Proteomic Data Analyses.* J Proteome Res **3**, 1002-1008 (2004).

73. Shinkawa, T., Taoka, M., Yamauchi, Y., Ichimura, T., Kaji, H., Takahashi, N. and Isobe, T. *Stem: A Software Tool for Large-Scale Proteomic Data Analyses.* J Proteome Res **4**, 1826-1831 (2005).
74. Hartler, J., Thallinger, G. G., Stocker, G., Sturn, A., Burkard, T. R., Körner, E., Rader, R., Schmidt, A., Mechtler, K. and Trajanoski, Z. *Maspectras: A Platform for Management and Analysis of Proteomics Lc-Ms/Ms Data.* BMC Bioinformatics **8**, 197-209 (2007).
75. Douglas, D. J., Frank, A. J. and Mao, D. *Linear Ion Traps in Mass Spectrometry.* Mass Spectrom Rev **24**, 1-29 (2005).
76. Scigelova, M. and Makarov, A. *Orbitrap Mass Analyzer - Overview and Applications in Proteomics.* Proteomics **6**, 16-21 (2006).
77. Marshall, A. G., Hendrickson, C. L. and Jackson, G. S. *Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer.* Mass Spectrom Rev **17**, 1-35 (1998).
78. Mikesh, L. M., Ueberheide, B., Chi, A., Coon, J. J., Syka, J. E. P., Shabanowitz, J. and Hunt, D. F. *The Utility of Etd Mass Spectrometry in Proteomic Analysis.* Biochim Biophys Acta **1764**, 1811-1822 (2006).
79. Cho, W. C. S. *Proteomics Technologies and Challenges.* Genomics Proteomics Bioinformatics **5**, 77-85 (2007).
80. Corthals, G. C., Wasinger, V. C., Hochstrasser , D. F. and Sanchez, J. C. *The Dynamic Range of Protein Expression: A Challenge for Proteomic Research.* Electrophoresis **21**, 1104-1115 (2000).
81. Anderson, N. L. and Anderson, N. G. *The Human Plasma Proteome: History, Character, and Diagnostic Prospects.* Mol Cell Proteomics **1**, 845-867 (2002).
82. Yamashita, M. and Fenn, J. B. *Electrospray Ion Source. Another Variation on the Free-Jet Theme.* J Phys Chem **88**, 4451-4459 (1984).
83. Yamashita, M. and Fenn, J. B. *Negative Ion Production with the Electrospray Ion Source.* J Phys Chem **88**, 4671-4675 (1984).
84. Karas, M. and Hillenkamp, F. in 11th International Mass Spectrometry Conference (Bordeaux, 1988).
85. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T. and Matsuo, T. *Protein and Polymer Analyses up to M/Z 100 000 by Laser Ionization Time-of-Flight Mass Spectrometry.* Rapid Commun Mass Spectrom **2**, 151-153 (1988).
86. Kolker, E., Higdon, R. and Hogan, J. M. *Protein Identification and Expression Analysis Using Mass Spectrometry.* Trends Microbiol **14**, 229-235 (2006).
87. Olsen, J. V., Ong, S.-E. and Mann, M. *Trypsin Cleaves Exclusively C-Terminal to Arginine and Lysine Residues.* Mol Cell Proteomics **3**, 608-614 (2004).
88. Wysocki, V. H., Resing, K. A., Zhang, Q. and Cheng, G. *Mass Spectrometry of Peptides and Proteins.* Methods **35**, 211-222 (2005).
89. Guerrera, I. and Kleiner, O. *Application of Mass Spectrometry in Proteomics.* Biosci Rep **25**, 71-93 (2005).
90. Calvo, K., Liotta, L. and Petricoin, E. *Clinical Proteomics: From Biomarker Discovery and Cell Signaling Profiles to Individualized Personal Therapy.* Biosci Rep **25**, 107-125 (2005).
91. Matt, P., Fu, Z., Fu, Q. and Van Eyk, J. E. *Biomarker Discovery: Proteome Fractionation and Separation in Biological Samples.* Physiol Genomics **in press (doi:10.1152/physiolgenomics.00282.2007)** (2007).

92. Veenstra, T. D. *Global and Targeted Quantitative Proteomics for Biomarker Discovery.* J Chromatogr B Analyt Technol Biomed Life Sci **847**, 3-11 (2007).
93. Ho, L., Sharma, N., Blackman, L., Festa, E., Reddy, G. and Pasinetti, G. M. *From Proteomics to Biomarker Discovery in Alzheimer's Disease.* Brain Res Rev **48**, 360-369 (2005).
94. Houtman, R. and van den Worm, E. *Asthma, the Ugly Duckling of Lung Disease Proteomics?* J Chromatogr B Biomed Sci Appl **815**, 285-294 (2005).
95. Xiao, Z., Prieto, D., Conrads, T. P., Veenstra, T. D. and Issaq, H. J. *Proteomic Patterns: Their Potential for Disease Diagnosis.* Mol Cell Endocrinol **230**, 95-106 (2005).
96. Rocken, C., Ebert, M. P. A. M. P. A. and Roessner, A. *Proteomics in Pathology, Research and Practice.* Pathol Res Pract **200**, 69-82 (2004).
97. Hanash, S. *Disease Proteomics.* Nature **422**, 226-232 (2003).
98. Zolg, J. W. and Langen, H. *How Industry Is Approaching the Search for New Diagnostic Markers and Biomarkers.* Mol Cell Proteomics **3**, 345-354 (2004).
99. Mathias, D. *Subcellular Proteomics.* Mass Spectrom Rev **22**, 27-56 (2003).
100. Huber, L. A., Pfaller, K. and Vietor, I. *Organelle Proteomics: Implications for Subcellular Fractionation in Proteomics.* Circ Res **92**, 962-968 (2003).
101. De Duve, C., Pressman, B. C., Gianetto, R., Wattiaux, R. and Appelmans, F. *Intracellular Distribution Patterns of Enzymes in Rat-Liver Tissue.* Biochemistry **60**, 604-617 (1955).
102. Cribb, A. E., Peyrou, M., Muruganandan, S. and Schneider, L. *The Endoplasmic Reticulum in Xenobiotic Toxicity.* Drug Metab Rev **37**, 405 - 442 (2005).
103. Sukhodub, A. L. and Burchell, A. *Preparation of Intact Microsomes from Cultured Mammalian H4iie Cells.* J Pharmacol Toxicol Methods **52**, 330-334 (2005).
104. Brunet, S., Thibault, P., Gagnon, E., Kearney, P., Bergeron, J. J. M. and Desjardins, M. *Organelle Proteomics: Looking at Less to See More.* Trends Cell Biol **13**, 629-638 (2003).
105. Andersen, J. S. and M., M. *Organellar Proteomics: Turning Inventories into Insights.* EMBO Rep **7**, 874–879 (2006).
106. Braun, R., Kinkl, N., Beer, M. and Ueffing, M. *Two-Dimensional Electrophoresis of Membrane Proteins.* Anal Bioanal Chem **389**, 1033-1045 (2007).
107. Yates Iii, J. R., Gilchrist, A., Howell, K. E. and Bergeron, J. J. M. *Proteomics of Organelles and Large Cellular Structures.* Nat Rev Mol Cell Biol **6**, 702-714 (2005).
108. Washburn, M. P., Wolters, D. and Yates, J. R. *Large-Scale Analysis of the Yeast Proteome by Multidimensional Protein Identification Technology.* Nat Biotech **19**, 242-247 (2001).
109. Delahunty, C. and III., J. R. Y. *Mudpit: Multidimensional Protein Identification Technology* BioTechniques **43**, 563-569 (2007).
110. Pan, S., Shi, M., Jin, J., Albin, R. L., Lieberman, A., Gearing, M., Lin, B., Pan, C., Yan, X., Kashima, D. T. and Zhang, J. *Proteomics Identification of Proteins in Human Cortex Using Multidimensional Separations and Maldi Tandem Mass Spectrometer.* Mol Cell Proteomics **6**, 1818-1823 (2007).
111. Kislinger, T., Gramolini, A. O., MacLennan, D. H. and Emili, A. *Multidimensional Protein Identification Technology (Mudpit): Technical Overview of a Profiling Method Optimized for the Comprehensive Proteomic Investigation of Normal and Diseased Heart Tissue.* J Am Soc Mass Spectrom **16**, 1207-1220 (2005).

112. Fujii, K., Nakano, T., Kawamura, T., Usui, F., Bando, Y., Wang, R. and Nishimura, T. *Multidimensional Protein Profiling Technology and Its Application to Human Plasma Proteome.* J Proteome Res **3**, 712-718 (2004).
113. Kopec, K. K., Bozyczko-Coyne, D. and Williams, M. *Target Identification and Validation in Drug Discovery: The Role of Proteomics.* Biochem Pharmacol **69**, 1133-1139 (2005).
114. De Meyts, P. *Insulin and Its Receptor: Structure, Function and Evolution.* Bioessays **26**, 1351-1362 (2004).
115. Meloni, B. P., Tilbrook, P. A., Boulos, S., Arthur, P. G. and Knuckey, N. W. *Erythropoietin Preconditioning in Neuronal Cultures: Signaling, Protection from in Vitro Ischemia, and Proteomic Analysis.* J Neurosci Res **83**, 584-593 (2006).
116. Carter, P., Smith, L. and Ryan, M. *Identification and Validation of Cell Surface Antigens for Antibody Targeting in Oncology.* Endocr Relat Cancer **11**, 659-687 (2004).
117. Ong, S.-E. and Mann, M. *Mass Spectrometry-Based Proteomics Turns Quantitative.* Nat Chem Biol **1**, 252-262 (2005).
118. Anderson, L. *Candidate-Based Proteomics in the Search for Biomarkers of Cardiovascular Disease.* J Physiol **563**, 23-60 (2005).
119. Sil, P., Mukherjee, D. and Sen, S. *Quantification of Myotrophin from Spontaneously Hypertensive and Normal Rat Hearts.* Circ Res **76**, 1020-1027 (1995).
120. Papoian, R., Duffy, F., Sanner, M. and Wilt, E. *A Sensitive Elisa for Measuring Recombinant Human Interleukin-3 in Human Plasma or Serum.* J Immunol Methods **145**, 161-165 (1991).
121. Kippen, A. D., Cerini, F., Vadas, L., Stocklin, R., Vu, L., Offord, R. E. and Rose, K. *Development of an Isotope Dilution Assay for Precise Determination of Insulin, C-Peptide, and Proinsulin Levels in Non-Diabetic and Type Ii Diabetic Individuals with Comparison to Immunoassay.* J Biol Chem **272**, 12513-12522 (1997).
122. Hamdan, M. and Righetti, P. G. *Modern Strategies for Protein Quantification in Proteome Analysis: Advantages and Limitations.* Mass Spectrom Rev **21**, 287-302 (2002).
123. Görg, A., Weiss, W. and Dunn, M. J. *Current Two-Dimensional Electrophoresis Technology for Proteomics.* Proteomics **4**, 3665-3685 (2004).
124. Marouga, R., David, S. and Hawkins, E. *The Development of the Dige System: 2d Fluorescence Difference Gel Analysis Technology.* Anal Bioanal Chem **382**, 669-678 (2005).
125. Gharbi, S., Gaffney, P., Yang, A., Zvelebil, M. J., Cramer, R., Waterfield, M. D. and Timms, J. F. *Evaluation of Two-Dimensional Differential Gel Electrophoresis for Proteomic Expression Analysis of a Model Breast Cancer Cell System.* Mol Cell Proteomics **1**, 91-98 (2002).
126. Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I. and Davison, M. *Validation and Development of Fluorescence Two-Dimensional Differential Gel Electrophoresis Proteomics Technology.* Proteomics **1**, 377-396 (2001).
127. Alban, A., David, S. O., Bjorkesten, L., Andersson, C., Sloge, E., Lewis, S. and Currie, I. *A Novel Experimental Design for Comparative Two-Dimensional Gel Analysis: Two-Dimensional Difference Gel Electrophoresis Incorporating a Pooled Internal Standard.* Proteomics **3**, 36-44 (2003).

128. Hoorn, E. J., Hoffert, J. D. and Knepper, M. A. *The Application of Dige-Based Proteomics to Renal Physiology.* Nephron Physiol **104**, 61-72 (2006).
129. Fenselau, C. *A Review of Quantitative Methods for Proteomic Studies.* J Chromatogr B **855**, 14-20 (2007).
130. Schneider, L. V. and Hall, M. P. *Stable Isotope Methods for High-Precision Proteomics.* Drug Discov Today **10**, 353-363 (2005).
131. Yan, W. and Chen, S. S. *Mass Spectrometry-Based Quantitative Proteomic Profiling.* Brief Funct Genomic Proteomic **4**, 27-38 (2005).
132. Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A. and Mann, M. *Stable Isotope Labeling by Amino Acids in Cell Culture, Silac, as a Simple and Accurate Approach to Expression Proteomics.* Mol Cell Proteomics **1**, 376-386 (2002).
133. Krijgsveld, J., Ketting, R. F., Mahmoudi, T., Johansen, J., Artal-Sanz, M., Verrijzer, C. P., Plasterk, R. H. A. and Heck, A. J. R. *Metabolic Labeling of C. Elegans and D. Melanogaster for Quantitative Proteomics.* Nat Biotech **21**, 927-931 (2003).
134. Ishihama, Y., Sato, T., Tabata, T., Miyamoto, N., Sagane, K., Nagasu, T. and Oda, Y. *Quantitative Mouse Brain Proteomics Using Culture-Derived Isotope Tags as Internal Standards.* Nat Biotech **23**, 617-621 (2005).
135. Munchbach, M., Quadroni, M., Miotto, G. and James, P. *Quantitation and Facilitated De Novo Sequencing of Proteins by Isotopic N-Terminal Labeling of Peptides with a Fragmentation-Directing Moiety.* Anal Chem **72**, 4047-4057 (2000).
136. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H. and Aebersold, R. *Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags.* Nat Biotech **17**, 994-999 (1999).
137. Oda, Y., Owa, T., Sato, T., Boucher, B., Daniels, S., Yamanaka, H., Shinohara, Y., Yokoi, A., Kuromitsu, J. and Nagasu, T. *Quantitative Chemical Proteomics for Identifying Candidate Drug Targets.* Anal Chem **75**, 2159-2165 (2003).
138. Zhou, H., Ranish, J. A., Watts, J. D. and Aebersold, R. *Quantitative Proteome Analysis by Solid-Phase Isotope Tagging and Mass Spectrometry.* Nat Biotech **20**, 512-515 (2002).
139. Zhou, H., Watts, J. D. and Aebersold, R. *A Systematic Approach to the Analysis of Protein Phosphorylation.* Nat Biotech **19**, 375-378 (2001).
140. Oda, Y., Nagasu, T. and Chait, B. T. *Enrichment Analysis of Phosphorylated Proteins as a Tool for Probing the Phosphoproteome.* Nat Biotech **19**, 379-382 (2001).
141. Liu, Y., Patricelli, M. P. and Cravatt, B. F. *Activity-Based Protein Profiling: The Serine Hydrolases.* Proc Natl Acad Sci U S A **96**, 14694-14699 (1999).
142. Zhang, H., Li, X.-j., Martin, D. B. and Aebersold, R. *Identification and Quantification of N-Linked Glycoproteins Using Hydrazide Chemistry, Stable Isotope Labeling and Mass Spectrometry.* Nat Biotech **21**, 660-666 (2003).
143. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A. and Pappin, D. J. *Multiplexed Protein Quantitation in Saccharomyces Cerevisiae Using Amine-Reactive Isobaric Tagging Reagents.* Mol Cell Proteomics **3**, 1154-1169 (2004).
144. Zieske, L. R. *A Perspective on the Use of Itraqtm Reagent Technology for Protein Complex and Profiling Studies.* J Exp Bot **57**, 1501-1508 (2006).

145. Pierce, A., Unwin, R. D., Evans, C. A., Griffiths, S., Carney, L., Zhang, L., Jaworska, E., Lee, C.-F., Blinco, D., Okoniewski, M. J., Miller, C. J., Bitton, D. A., Spooncer, E. and Whetton, A. D. *Eight-Channel Itraq Enables Comparison of the Activity of 6 Leukaemogenic Tyrosine Kinases.* Mol Cell Proteomics, M700251-MCP700200 (2007).

146. Schnölzer, M., Jedrzejewski, P. and Lehmann, W. D. *Protease-Catalyzed Incorporation of 18o into Peptide Fragments and Its Application for Protein Sequencing by Electrospray and Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry.* Electrophoresis **17**, 945-953 (1996).

147. Yao, X., Freas, A., Ramirez, J., Demirev, P. A. and Fenselau, C. *Proteolytic 18o Labeling for Comparative Proteomics: Model Studies with Two Serotypes of Adenovirus.* Anal Chem **73**, 2836-2842 (2001).

148. Yao, X., Afonso, C. and Fenselau, C. *Dissection of Proteolytic 18o Labeling: Endoprotease-Catalyzed 16o-to-18o Exchange of Truncated Peptide Substrates.* J Proteome Res **2**, 147-152 (2003).

149. Miyagi, M. and Bao, K. C. S. *Proteolytic 18o-Labeling Strategies for Quantitative Proteomics.* Mass Spectrom Rev **26**, 121-136 (2007).

150. Ramos-Fernandez, A., Lopez-Ferrer, D. and Vazquez, J. *Improved Method for Differential Expression Proteomics Using Trypsin-Catalyzed 18o Labeling with a Correction for Labeling Efficiency.* Mol Cell Proteomics **6**, 1274-1286 (2007).

151. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. and Kuster, B. *Quantitative Mass Spectrometry in Proteomics: A Critical Review.* Anal Bioanal Chem **389**, 1017-1031 (2007).

152. Kirkpatrick, D. S., Gerber, S. A. and Gygi, S. P. *The Absolute Quantification Strategy: A General Procedure for the Quantification of Proteins and Post-Translational Modifications.* Methods **35**, 265-273 (2005).

153. Desiderio, D. M., Kai, M., Tanzer, F. S., Trimble, J. and Wakelyn, C. *Measurement of Enkephalin Peptides in Canine Brain Regions, Teeth, and Cerebrospinal Fluid with High-Performance Liquid Chromatography and Mass Spectrometry.* J Chromatogr A **297**, 245-260 (1984).

154. Barr, J. R., Maggio, V. L., Patterson, D. G., Jr., Cooper, G. R., Henderson, L. O., Turner, W. E., Smith, S. J., Hannon, W. H., Needham, L. L. and Sampson, E. J. *Isotope Dilution-Mass Spectrometric Quantification of Specific Proteins: Model Application with Apolipoprotein a-I.* Clin Chem **42**, 1676-1682 (1996).

155. Barnidge, D. R., Dratz, E. A., Martin, T., Bonilla, L. E., Moran, L. B. and Lindall, A. *Absolute Quantification of the G Protein-Coupled Receptor Rhodopsin by Lc/Ms/Ms Using Proteolysis Product Peptides and Synthetic Peptide Standards.* Anal Chem **75**, 445-451 (2003).

156. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W. and Gygi, S. P. *Absolute Quantification of Proteins and Phosphoproteins from Cell Lysates by Tandem Ms.* Proc Natl Acad Sci U S A **100**, 6940-6945 (2003).

157. Kondrat, R. W., McClusky, G. A. and Cooks, R. G. *Multiple Reaction Monitoring in Mass Spectrometry/Mass Spectrometry for Direct Analysis of Complex Mixtures.* Anal Chem **50**, 2017-2021 (1978).

158. Ong, S.-E., Foster, L. J. and Mann, M. *Mass Spectrometric-Based Approaches in Quantitative Proteomics.* Methods **29**, 124-130 (2003).

159. Stemmann, O., Zou, H., Gerber, S. A., Gygi, S. P. and Kirschner, M. W. *Dual Inhibition of Sister Chromatid Separation at Metaphase.* Cell **107**, 715-726 (2001).
160. Nakajima, T. and Aoyama, T. *Polymorphism of Drug-Metabolizing Enzymes in Relation to Individual Susceptibility to Industrial Chemicals.* Ind Health **38**, 143-152 (2000).
161. Iyer, K. R. and Sinz, M. W. *Characterization of Phase I and Phase Ii Hepatic Drug Metabolism Activities in a Panel of Human Liver Preparations.* Chem Biol Interact **118**, 151-169 (1999).
162. Zhang, J. Y., Fen Wang, Y. and Prakash, C. *Xenobiotic-Metabolizing Enzymes in Human Lung.* Curr Drug Metab **7**, 939-948 (2006).
163. Miller, D. S. *Xenobiotic Export Pumps, Endothelin Signaling, and Tubular Nephrotoxicants - a Case of Molecular Hijacking.* J Biochem Mol Toxicol **16**, 121-127 (2002).
164. Raunio, H., Husgafvel-Pursiainen, K., Anttila, S., Hietanen, E., Hirvonen, A. and Pelkonen, O. *Diagnosis of Polymorphisms in Carcinogen-Activating and Inactivating Enzymes and Cancer Susceptibility-a Review.* Gene **159**, 113-121 (1995).
165. Smith, G., Wolf, C. R., Deeni, Y. Y., Dawe, R. S., Evans, A. T., Comrie, M. M., Ferguson, J. and Ibbotson, S. H. *Cutaneous Expression of Cytochrome P450 Cyp2s1: Individuality in Regulation by Therapeutic Agents for Psoriasis and Other Skin Diseases.* Lancet **361**, 1336-1343 (2003).
166. Dorne, J. L. C. M., Walton, K. and Renwick, A. G. *Polymorphic Cyp2c19 and N-Acetylation: Human Variability in Kinetics and Pathway-Related Uncertainty Factors.* Food Chem Toxicol **41**, 225-245 (2003).
167. Norppa, H. *Genetic Susceptibility, Biomarker Respones, and Cancer.* Mutat Res **544**, 339-348 (2003).
168. Vineis, P. *The Relationship between Polymorphisms of Xenobiotic Metabolizing Enzymes and Susceptibility to Cancer.* Toxicology **181-182**, 457-462 (2002).
169. Alaejos, M. S., Pino, V. and Afonso, A. M. *Metabolism and Toxicology of Heterocyclic Aromatic Amines When Consumed in Diet: Influence of the Genetic Susceptibility to Develop Human Cancer. A Review.* Food Research International **In Press** (2008).
170. Kiyohara, C., Otsu, A., Shirakawa, T., Fukuda, S. and Hopkin, J. M. *Genetic Polymorphisms and Lung Cancer Susceptibility: A Review.* Lung Cancer **37**, 241-256 (2002).
171. Kiani, J. and Imam, S. Z. *Medicinal Importance of Grapefruit Juice and Its Interaction with Various Drugs.* Nutr J **6**, 33-41 (2007).
172. Gochfeld, M. *Framework for Gender Differences in Human and Animal Toxicology.* Environ Res **104**, 4-21 (2007).
173. Hukkanen, J., Jacob, P., III and Benowitz, N. L. *Metabolism and Disposition Kinetics of Nicotine.* Pharmacol Rev **57**, 79-115 (2005).
174. Nagle, C. M., Chenevix-Trench, G., Spurdle, A. B. and Webb, P. M. *The Role of Glutathione-S-Transferase Polymorphisms in Ovarian Cancer Survival.* Eur J Cancer **43**, 283-290 (2007).
175. Patterson, L. H. and Murray, G. I. *Tumour Cytochrome P450 and Drug Activation.* Curr Pharm Des **8**, 1335-1347 (2002).
176. Purnapatre, K., Khattar, S. K. and Saini, K. S. *Cytochrome P450s in the Development of Target-Based Anticancer Drugs.* Cancer Lett **259**, 1-15 (2008).

177. Omura, T. and Sato, R. *A New Cytochrome in Liver Microsomes.* J Biol Chem **237**, 1375-1376 (1962).
178. Estabrook, R. W. *The Remarkable P450s: A Historical Overview of These Versatile Hemeprotein Catalysts.* FASEB J **10**, 202-204 (1996).
179. Nebert, D. W. and Gonzalez, F. J. *P450 Genes: Structure, Evolution, and Regulation.* Annu Rev Biochem **56**, 945-993 (1987).
180. Nelson, D. R. and Strobel, H. W. *Evolution of Cytochrome P-450 Proteins.* Mol Biol Evol **4**, 572-593 (1987).
181. Guengerich, F. P. *Cytochrome P450: What Have We Learned and What Are the Future Issues?* Drug Metab Rev **36**, 159 - 197 (2004).
182. Seliskar, M. and Rozman, D. *Mammalian Cytochromes P450--Importance of Tissue Specificity.* Biochim Biophys Acta **1770**, 458-466 (2007).
183. Rendic, S. and Di Carlo, F. J. *Human Cytochrome P450 Enzymes: A Status Report Summarizing Their Reactions, Substrates, Inducers, and Inhibitors.* Drug Metab Rev **29**, 413-580 (1997).
184. Omura, T. *Mitochondrial P450s.* Chem Biol Interact **163**, 86-93 (2006).
185. Richter, T., Murdter, T. E., Heinkele, G., Pleiss, J., Tatzel, S., Schwab, M., Eichelbaum, M. and Zanger, U. M. *Potent Mechanism-Based Inhibition of Human Cyp2b6 by Clopidogrel and Ticlopidine.* J Pharmacol Exp Ther **308**, 189-197 (2004).
186. Guengerich, F. P. *Cytochromes P450, Drugs, and Diseases.* Mol Interv **3**, 194-204 (2003).
187. Lewis, D. F. V. *On the Recognition of Mammalian Microsomal Cytochrome P450 Substrates and Their Characteristics: Towards the Prediction of Human P450 Substrate Specificity and Metabolism.* Biochem Pharmacol **60**, 293-306 (2000).
188. Gillam, E. M. J. *Engineering Cytochrome P450 Enzymes.* Chem Res Tox **21**, 220-231 (2008).
189. Sansen, S., Yano, J. K., Reynald, R. L., Schoch, G. A., Griffin, K. J., Stout, C. D. and Johnson, E. F. *Adaptations for the Oxidation of Polycyclic Aromatic Hydrocarbons Exhibited by the Structure of Human P450 1a2.* J Biol Chem **282**, 14348-14355 (2007).
190. Smith, B. D., Sanders, J. L., Porubsky, P. R., Lushington, G. H., Stout, C. D. and Scott, E. E. *Structure of the Human Lung Cytochrome P450 2a13.* J Biol Chem **282**, 17306-17313 (2007).
191. Williams, P. A., Cosme, J., Vinkovic, D. M., Ward, A., Angove, H. C., Day, P. J., Vonrhein, C., Tickle, I. J. and Jhoti, H. *Crystal Structures of Human Cytochrome P450 3a4 Bound to Metyrapone and Progesterone.* Science **305**, 683-686 (2004).
192. Williams, P. A., Cosme, J., Ward, A., Angove, H. C., Matak Vinkovic, D. and Jhoti, H. *Crystal Structure of Human Cytochrome P450 2c9 with Bound Warfarin.* Nature **424**, 464-468 (2003).
193. Schoch, G. A., Yano, J. K., Wester, M. R., Griffin, K. J., Stout, C. D. and Johnson, E. F. *Structure of Human Microsomal Cytochrome P450 2c8: Evidence for a Peripheral Fatty Acid Binding Site.* J Biol Chem **279**, 9497-9503 (2004).
194. Yano, J. K., Hsu, M.-H., Griffin, K. J., Stout, C. D. and Johnson, E. F. *Structures of Human Microsomal Cytochrome P450 2a6 Complexed with Coumarin and Methoxsalen.* Nat Struct Mol Biol **12**, 822-823 (2005).
195. Rowland, P., Blaney, F. E., Smyth, M. G., Jones, J. J., Leydon, V. R., Oxbrow, A. K., Lewis, C. J., Tennant, M. G., Modi, S., Eggleston, D. S., Chenery, R. J.

and Bridges, A. M. *Crystal Structure of Human Cytochrome P450 2d6.* J Biol Chem **281**, 7614-7622 (2006).

196. Danielson, P. B. *The Cytochrome P450 Superfamily: Biochemistry, Evolution and Drug Metabolism in Humans.* Curr Drug Metab **3**, 561-597 (2002).
197. Hlavica, P. and Lewis, D. F. V. *Allosteric Phenomena in Cytochrome P450-Catalyzed Monooxygenations.* Eur J Biochem **268**, 4817-4832 (2001).
198. Guengerich, F. P. *Mechanisms of Cytochrome P450 Substrate Oxidation: Minireview.* J Biochem Mol Toxicol **21**, 163-168 (2007).
199. Ortiz de Montellano, P. R. and De Voss, J. J. *Oxidizing Species in the Mechanism of Cytochrome P450.* Nat Prod Rep **19**, 477-493 (2002).
200. Lane, C. S. The Analysis of Cytochrome P450 Proteins by Mass Spectrometry (Thesis) (University of London, London, 2004).
201. Baris, N., Kalkan, S., Güneri, S., Bozdemir, V. and Guven, H. *Influence of Carvedilol on Serum Digoxin Levels in Heart Failure: Is There Any Gender Difference?* Eur J Clin Pharmacol **62**, 535-538 (2006).
202. Patel, H., Bell, D., Molokhia, M., Srishanmuganathan, J., Patel, M., Car, J. and Majeed, A. *Trends in Hospital Admissions for Adverse Drug Reactions in England: Analysis of National Hospital Episode Statistics 1998–2005.* BMC Clin Pharmacology **7** (2007).
203. Abu-Abed, S., Dolle, P., Metzger, D., Beckett, B., Chambon, P. and Petkovich, M. *The Retinoic Acid-Metabolizing Enzyme, Cyp26a1, Is Essential for Normal Hindbrain Patterning, Vertebral Identity, and Development of Posterior Structures.* Genes Dev **15**, 226-240 (2001).
204. White, P. C. *Aldosterone Synthase Deficiency and Related Disorders.* Mol Cell Endocrinol **217**, 81-87 (2004).
205. Guengerich, F. P. *Cytochrome P450, Drugs, and Diseases.* Mol Interv **3**, 194-204 (2003).
206. Guengerich, F. P. *Cytochrome P450s and Other Enzymes in Drug Metabolism and Toxicity.* AAPS J **8**, E101-E111 (2006).
207. Koop, D. R. *Oxidative and Reductive Metabolism by Cytochrome P450 2e1.* FASEB J **6**, 724-730 (1992).
208. Tanaka, E., Terada, M. and Misawa, S. *Cytochrome P450 2e1:Its Clinical and Toxicological Role.* J Clin Pharm Ther **15**, 165-175 (2000).
209. Danko, I. M. and Chaschin, N. A. *Association of Cyp2e1 Gene Polymorphism with Predisposition to Cancer Development.* Exp Oncol **27**, 248-256 (2005).
210. Bruno, R. D. and Njar, V. C. O. *Targeting Cytochrome P450 Enzymes: A New Approach in Anti-Cancer Drug Development.* Bioorg Med Chem **15**, 5047-5060 (2007).
211. Haufroid, V., Toubeau, F., Clippe, A., Buyschaert, M., Gala, J. L. and Lison, D. *Real-Time Quantification of Cytochrome P4502e1mrna in Human Peripheral Blood Lymphocytes by Reverse Transcription-Pcr:Method and Practical Application.* Clin Chem **47**, 1126-1127 (2001).
212. Dey, A., Kessova, I. G. and Cederbaum, A. I. *Decreased Protein and Mrna Expression of Er Stress Proteins Grp78 and Grp94 in Hepg2 Cells over-Expressing Cyp2e1.* Arch Biochem Biophys **447**, 155-166 (2006).
213. Tanaka, E., Kurata, N. and H., Y. *How Useful Is the Cocktail Approach for Evaluating Human Hepatic Drug Metabolizing Capacity Using Cytochrome P450 Phenotyping Probes in Vivo.* J Clin Pharm Ther **28**, 157-165 (2003).

214. Wang, Y., Al-Gazzar, A., Seibert, C., Sharif, A., Lane, C. and Griffiths, W. J. *Proteomic Analysis of Cytochromes P450: A Mass Spectrometry Approach.* Biochem Soc Trans **34**, 1246-1251 (2006).
215. Guengerich, F. P. *Cytochrome P-450 3a4: Regulation and Role in Drug Metabolism.* Annu Rev Pharmacol Toxicol **39**, 1-17 (1999).
216. Rodrigues, A. D. *Integrated Cytochrome P450 Reaction Phenotyping.* Biochem Pharmacol **57**, 465-480 (1999).
217. Shimada, T., Yamazaki, H., Mimura, M., Inui, Y. and Guengerich, F. P. *Interindividual Variations in Human Liver Cytochrome P-450 Enzymes Involved in the Oxidation of Drugs, Carcinogens and Toxic Chemicals:Studies with Liver Microsomes of 30 Japanese and 30 Caucasians.* J Pharmacol Exp Ther **270**, 414-423 (1994).
218. Cederbaum, A. I. *Cyp2e1-Biochemical and Toxicological Aspects and Role in Alcohol-Induced Liver Injury.* Mt Sinai J Med **73**, 657-672 (2006).
219. Lieber, C. S. *Ethanol Metabolism, Cirrhosis and Alcoholism.* Clin Chim Acta **257**, 59-84 (1997).
220. Winters, D. K. and Cederbaum, A. I. *The Content and Activity of Cytochrome P450 2e1 in Liver Microsomes from Alcohol-Preferring and Non-Peferring Rats.* Alcohol Alcohol **27**, 63-70 (1992).
221. Gonzalez, F. J. *Role of Cytochromes P450 in Chemical Toxicity and Oxidative Stress: Studies with Cyp2e1.* Mutat Res **569**, 101-110 (2005).
222. Gonzalez, F. J. *The 2006 Bernard B. Brodie Award Lecture. Cyp2e1.* Drug Metab Dispos **35**, 1-8 (2007).
223. Guengerich, F. P., Shimada, T., Yun, C. H., Yamazaki, Y., Raney, K. D., Thier, R., Coles, B. and Harris, T. M. *Interactions of Ingested Food, Beverage, and Tobacco Components Involving Human Cytochrome P4501a2, 2a6, 2e1, and 3a4 Enzymes.* Environ Health Perspect **102**, 49-53 (1994).
224. Diamond, D. L., Proll, S. C., Jacobs, J. M., Chan, E. Y., Camp, D. G. I., Smith, R. D. and Katze, M. G. *Hepatoproteomics: Applying Proteomic Technologies to the Study of Liver Function and Disease.* Hepatology **44**, 299-308 (2006).
225. Sun, S., Lee, N. P. Y., Poon, R. T. P., Fan, S.-T., He, Q. Y., Lau, G. K. and Luk, J. M. *Oncoproteomics of Hepatocellular Carcinoma: From Cancer Markers' Discovery to Functional Pathways.* Liver Int **27**, 1021-1038 (2007).
226. Diamond, D. L., Jacobs, J. M., Paeper, B., Proll, S. C., Gritsenko, M. A., Carithers, R. L., Jr., Larson, A. M., Yeh, M. M., Camp, D. G., 2nd, Smith, R. D. and Katze, M. G. *Proteomic Profiling of Human Liver Biopsies: Hepatitis C Virus-Induced Fibrosis and Mitochondrial Dysfunction.* Hepatology **46**, 649-657 (2007).
227. Nabetani, T., Tabuse, Y., Tsugita, A. and Shoda, J. *Proteomic Analysis of Livers of Patients with Primary Hepatolithiasis.* Proteomics **5**, 1043-1061 (2005).
228. Simpson, R. J., Bernard, O. K., Greening, D. W. and Moritz, R. L. *Proteomics-Driven Cancer Biomarker Next Term Discovery: Looking to the Future* Curr Opin Chem Biol **12**, 72-77 (2008).
229. Noel-Georis, I., Bernard, A., Falmagne, P. and Wattiez, R. *Database of Bronchoalveolar Lavage Fluid Proteins.* J Chromatogr B Analyt Technol Biomed Life Sci **771**, 221-236 (2002).
230. Ostrowski, L. E., Blackburn, K., Radde, K. M., Moyer, M. B., Schlatzer, D. M., Moseley, A. and Boucher, R. C. *A Proteomic Analysis of Human Cilia: Identification of Novel Components.* Mol Cell Proteomics **1**, 451-465 (2002).

231. Li, C., Zhan, X., Li, M., Wu, X., Li, F., Li, J., Xiao, Z., Chen, Z., Feng, X., Chen, P., Xie, J. and Liang, S. *Proteomic Comparison of Two-Dimensional Gel Electrophoresis Profiles from Human Lung Squamous Carcinoma and Normal Bronchial Epithelial Tissues.* Genomics Proteomics Bioinformatics **1**, 58-67 (2003).
232. Alaiya, A. A., Roblick, U. J., Franzen, B., Bruch, H.-P. and Auer, G. *Protein Expression Profiling in Human Lung, Breast, Bladder, Renal, Colorectal and Ovarian Cancers.* J Chromatogr B **787**, 207-222 (2003).
233. Hukkanen, J. Xenobiotic-Metanolizing Cytochrome P450 Enzymes in Human Lung (Thesis) (University of Oulu, Oulu, 2000).
234. Bernauer, U., Heinrich-Hirsch, B., Tonnies, M., Peter-Matthias, W. and Gundert-Remy, U. *Characterisation of the Xenobiotic-Metabolizing Cytochrome P450 Expression Pattern in Human Lung Tissue by Immunochemical and Activity Determination.* Toxicol Lett **164**, 278-288 (2006).
235. Shimada, T., Yun, C. H., Yamazaki, H., Gautier, J. C., Beaune, P. H. and Guengerich, F. P. *Characterization of Human Lung Microsomal Cytochrome P-450 1a1 and Its Role in the Oxidation of Chemical Carcinogens.* Mol Pharmacol **41**, 856-864 (1992).
236. Hukkanen, J., Pelkonen, O. and Raunio, H. *Expression of Xenobiotic-Metabolizing Enzymes in Human Pulmonary Tissue: Possible Role in Susceptibility for Ild.* Eur Respir J **18**, 122S-126 (2001).
237. Bradford, M. M. *A Rapid and Sensitive Method for the Quantification of Microgram Quantities of Protein Utilizing the Principle of Dyebinding.* Anal Biochem **72**, 248-254 (1976).
238. Laemmli, U. K. *Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4.* Nature **227**, 680-685 (1970).
239. Savitzky, A. and Golay, M. J. E. *Smoothing and Differentiation of Data by Simplified Least Squares Procedures.* Anal Chem **36**, 1627-1639 (1964).
240. Kersey, P. J., Duarte, J., Williams, A., Karavidopoulou, Y., Birney, E. and Apweiler, R. *The International Protein Index: An Integrated Database for Proteomics Experiments.* Proteomics **4**, 1985-1988 (2004).
241. Ihaka, R. and Gentleman, R. *R: A Language for Data Analysis and Graphics.* J Comp Graph Statist **5**, 299-314 (1996).
242. Bouchardy, C., Benhamou, S., Jourenkova, N., Dayer, P. and Hirvonen, A. *Metabolic Genetic Polymorphisms and Susceptibility to Lung Cancer.* Lung Cancer **32**, 109-112 (2001).
243. Lane, C. S., Nisar, S., Griffiths, W. J., Davidson, B. R., Hewes, J., Welham, K. J. and Patterson, L. H. *Identification of Cytochrome P450 Enzymes in Human Colorectal Metastases and the Surrounding Liver.* Eur J Cancer **40**, 2127-2134 (2004).
244. Caro, A. A. and Cederbaum, A. I. *Oxidative Stress, Toxicology and Pharmacology of Cyp2e1.* Annu Rev Pharmacol Toxicol **44**, 27-42 (2004).
245. Nisar, S. Nano-Scale Liquid Chromatography Electrospray Tandem Mass Spectrometric Identification of Cytochrome P450 Isoforms in Tissues (Thesis) (University of London, London, 2005).
246. McManus, M. E., Boobis, A. R., Pacifici, G. M., Frempong, R. Y., Brodie, M. J., Kahn, G. C., Whyte, C. and Davies, D. S. *Xenobiotic Metabolism in the Human Lung.* Life Sci **26**, 481-487 (1980).

247. Wheeler, C. W. and Guenthner, T. M. *Spectroscopic Quantitation of Cytochrome P-450 in Human Lung Microsomes.* J Biochem Toxicol **5**, 269-272 (1990).
248. Kim, J. H., Sherman, M. E., Curriero, F. C., Guengerich, F. P., Strickland, P. T. and Sutter, T. R. *Expression of Cytochromes P450 1a1 and 1b1 in Human Lung from Smokers, Non-Smokers, and Ex-Smokers.* Toxicol Appl Pharmacol **199**, 210-219 (2004).
249. Serabjit-Singh, C. J., Wolf, C. R., Philpot, R. M. and Plopper, C. G. *Cytochrome P-450: Localization in Rabbit Lung.* Science **207**, 1469-1470 (1980).
250. Minchin, R. F. and Boyd, M. R. *Localization of Metabolic Activation and Deactivation Systems in the Lung: Significance to the Pulmonary Toxicity of Xenobiotics.* Annu Rev Pharmacol Toxicol **23**, 217-238 (1983).
251. Nisar, S., Lane, C. S., Wilderspin, A. F., Welham, K. J., Griffiths, W. J. and Patterson, L. H. *A Proteomic Approach to the Identification of Cytochrome P450 Isoforms in Male and Female Rat Liver by Nanoscale Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry.* Drug Metab Dispos **32**, 382-386 (2004).
252. Gygi, S. P. and Aebersold, R. *Mass Spectrometry and Proteomics.* Curr Opin Chem Biol **4**, 489-494 (2000).
253. Righetti, P. G., Campostrini, N., Pascali, J., Hamdanb, M. and Astner, H. *Quantitative Proteomics: A Review of Different Methodologies.* Eur J Mass Spectrom **10**, 335-348 (2004).
254. U. S. Department of Health, F. a. D. A. (Rockville, 2001).
255. Petushkova, N. A., Kanaeva, I. P., Lisitsa, A. V., Sheremetyeva, G. F., Zgoda, V. G., Samenkova, N. F., Karuzina, I. I. and Archakov, A. I. *Characterization of Human Liver Cytochromes P450 by Combining the Biochemical and Proteomic Approaches.* Toxicol in Vitro **20**, 966-974 (2006).
256. Afsar, A., Lee, C. and Riddick, D. S. *Modulation of the Expression of Constitutive Rat Hepatic Cytochrome P450 Isozymes by 5-Fluorouracil.* Can J Physiol Pharmacol **74**, 150-156 (1996).
257. Jenkins, R. E., Kitteringham, N. R., Hunter, C. L., Webb, S., Hunt, T. J., Elsby, R., Watson, R. B., Williams, D., Pennington, S. R. and Park, B. K. *Relative and Absolute Quantitative Expression Profiling of Cytochromes P450 Using Isotope-Coded Affinity Tags.* Proteomics **6**, 1934-1947 (2006).
258. Landi, M. T., Sinha, R., Lang, N. P. and Kadlubar, F. F. *Human Cytochrome P4501a2.* IARC Sci Publ **148**, 173-195 (1999).
259. Guengerich, F. P. and Turvy, C. G. *Comparison of Levels of Several Human Microsomal Cytochrome P-450 Enzymes and Epoxide Hydrolase in Normal and Disease States Using Immunochemical Analysis of Surgical Liver Samples.* J Pharmacol Exp Ther **256**, 1189-1194 (1991).
260. Stone, K. L. and WIlliams, K. R. in The Protein Protocols Handbook (ed. Walker, J. M.) 415-425 (Humana Press Inc., Totowa, 1996).
261. Ericsson, C., Franzen, B. and Nister, M. *Frozen Tissue Biobanks. Tissue Handling, Cryopreservation, Extraction, and Use for Proteomic Analysis.* Acta Oncol **45**, 643 - 661 (2006).
262. Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y. and Aebersold, R. *Evaluation of Two-Dimensional Gel Electrophoresis-Based Proteome Analysis Technology.* Proc Natl Acad Sci U S A **97**, 9390-9395 (2000).

263. Kislinger, T., Rahman, K., Radulovic, D., Cox, B., Rossant, J. and Emili, A. *Prism, a Generic Large Scale Proteomic Investigation Strategy for Mammals.* Mol Cell Proteomics **2**, 96-106 (2003).
264. Liu, H., Sadygov, R. G. and Yates, J. R. *A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics.* Anal Chem **76**, 4193-4201 (2004).
265. Chen, H. s., Rejtar, T., Andreev, V., Moskovets, E. and Karger, B. L. *Enhanced Characterization of Complex Proteomic Samples Using Lc-Maldi Ms/Ms: Exclusion of Redundant Peptides from Ms/Ms Analysis in Replicate Runs.* Anal Chem **77**, 7816-7825 (2005).
266. Linke, T., Doraiswamy, S. and Harrison, E. H. *Rat Plasma Proteomics: Effects of Abundant Protein Depletion on Proteomic Analysis.* J Chromatogr B **849**, 273-281 (2007).
267. Issaq, H. J., Xiao, Z. and Veenstra, T. D. *Serum and Plasma Proteomics.* Chem Rev **107**, 3601-3620 (2007).
268. Granger, J., Siddiqui, J., Copeland, S. and Remick, D. *Albumin Depletion of Human Plasma Also Removes Low Abundance Proteins Including the Cytokines.* Proteomics **5**, 4713-4718 (2005).
269. Guengerich, F. P., Dannan, G. A., Wright, S. T., Martin, M. V. and Kaminsky, L. S. *Purification and Characterization of Liver Microsomal Cytochromes P-450: Electrophoretic, Spectral, Catalytic, and Immunochemical Properties and Inducibility of Eight Isozymes Isolated from Rats Treated with Phenobarbital Or .Beta.-Naphthoflavone.* Biochemistry **21**, 6019-6030 (1982).
270. Jacobs, J. M., Adkins, J. N., Qian, W. J., Liu, T., Shen, Y., Camp, D. G. and Smith, R. D. *Utilizing Human Blood Plasma for Proteomic Biomarker Discovery.* J Proteome Res **4**, 1073-1085 (2005).
271. Song, Y., Hao, Y., Sun, A., Li, T., Li, W., Guo, L., Yan, Y., Geng, C., Chen, N., Zhong, F., Wei, H., Jiang, Y. and He, F. *Sample Preparation Project for the Subcellular Proteome of Mouse Liver.* Proteomics **6**, 5269-5277 (2006).
272. Oinonen, T. and Lindros, K. O. *Zonation of Hepatic Cytochrome P-450 Expression and Regulation.* Biochem J **329**, 17-35 (1998).
273. Edwards, R. J., Boobis, A. R. and Davies, D. S. *A Strategy for Investigating the Cyp Superfamily Using Targeted Antibodies Is a Paradigm for Functional Genomic Studies.* Drug Metab Dispos **31**, 1476-1480 (2003).
274. Kornilayev, B. A. and Alterman, M. A. *Utility of Polyclonal Antibodies Targeted toward Unique Tryptic Peptides in the Proteomic Analysis of Cytochrome P450 Isozymes.* Toxicology in Vitro **In Press** (2008).
275. Alterman, M. A., Kornilayev, B., Duzhak, T. and Yakovlev, D. *Quantitative Analysis of Cytochrome P450 Isozymes by Means of Unique Isozyme-Specific Tryptic Peptides: A Proteomic Approach.* Drug Metab Dispos **33**, 1399-1407 (2005).
276. Aebersold, R. *Constellations in a Cellular Universe.* Nature **422**, 115-116 (2003).
277. Speicher, K. D., Kolbas, O., Harper, S. and Speicher, D. W. *Systematic Analysis of Peptide Recoveries from in-Gel Digestions for Protein Identifications in Proteome Studies.* J Biomol Tech **11**, 74-86 (2000).
278. Barnidge, D. R., Hall, G. D., Stocker, J. L. and Muddiman, D. C. *Evaluation of a Cleavable Stable Isotope Labeled Synthetic Peptide for Absolute Protein Quantification Using Lc-Ms/Ms.* J Proteome Res **3**, 658-661 (2004).

279. Beynon, R. J., Doherty, M. K., Pratt, J. M. and Gaskell, S. J. *Multiplexed Absolute Quantification in Proteomics Using Artificial Qcat Proteins of Concatenated Signature Peptides.* Nat Meth **2**, 587-589 (2005).

280. Ji, Q. C., Rodila, R., Gage, E. M. and El-Shourbagy, T. A. *A Strategy of Plasma Protein Quantitation by Selective Reaction Monitoring of an Intact Protein.* Anal Chem **75**, 7008-7014 (2003).

281. Becher, F., Pruvost, A., Clement, G., Tabet, J. C. and Ezan, E. *Quantification of Small Therapeutic Proteins in Plasma by Liquid Chromatography-Tandem Mass Spectrometry: Application to an Elastase Inhibitor Epi-Hne4.* Anal Chem **78**, 2306-2313 (2006).

282. Bunk, D. M. and Welch, M. J. *Electrospray Ionization Mass Spectrometry for the Quantitation of Albumin in Human Serum.* J Am Soc Mass Spectrom **8**, 1247-1254 (1997).

283. Ji, Q. C., Gage, E. M., Rodila, R., Chang, M. S. and El-Shourbagy, T. A. *Method Development for the Concentration Determination of a Protein in Human Plasma Utilizing 96-Well Solid-Phase Extraction and Liquid Chromatography/Tandem Mass Spectrometric Detection.* Rapid Commun Mass Spectrom **17**, 794-799 (2003).

284. Duan, X., Chen, X., Yang, Y. and Zhong, D. *Precolumn Derivatization of Cysteine Residues for Quantitative Analysis of Five Major Cytochrome P450 Isoenzymes by Liquid Chromatography/Tandem Mass Spectrometry.* Rapid Commun Mass Spectrom **21**, 3234-3244 (2007).

285. Listgarten, J. and Emili, A. *Statistical and Computational Methods for Comparative Proteomic Profiling Using Liquid Chromatography-Tandem Mass Spectrometry.* Mol Cell Proteomics **4**, 419-434 (2005).

286. Mueller, L. N., Brusniak, M.-Y., Mani, D. R. and Aebersold, R. *An Assessment of Software Solutions for the Analysis of Mass Spectrometry Based Quantitative Proteomics Data.* J Proteome Res **7**, 51-61 (2008).

287. Tabata, T., Sato, T., Kuromitsu, J. and Oda, Y. *Pseudo Internal Standard Approach for Label-Free Quantitative Proteomics.* Anal Chem **79**, 8440-8445 (2007).

288. Bakhtiar, R. and Majumdar, T. K. *Tracking Problems and Possible Solutions in the Quantitative Determination of Small Molecule Drugs and Metabolites in Biological Fluids Using Liquid Chromatography-Mass Spectrometry.* J Pharmacol Toxicol Methods **55**, 227-243 (2007).

289. Decaestecker, T. N., Clauwaert, K. M., Van Bocxlaer, J. F., Lambert, W. E., Van den Eeckhout, E. G., Van Peteghem, C. H. and De Leenheer, A. P. *Evaluation of Automated Single Mass Spectrometry to Tandem Mass Spectrometry Function Switching for Comprehensive Drug Profiling Analysis Using a Quadrupole Time-of-Flight Mass Spectrometer.* Rapid Commun Mass Spectrom **14**, 1787-1792 (2000).

290. James, P. F., Perugini, M. A. and O'Hair, R. A. J. *Sources of Artefacts in the Electrospray Ionization Mass Spectra of Saturated Diacylglycerophosphocholines: From Condensed Phase Hydrolysis Reactions through to Gas Phase Intercluster Reactions.* J Am Soc Mass Spectrom **17**, 384-394 (2006).

291. Damen, C. W. N., Rosing, H., Schellens, J. H. M. and Beijnen, J. H. *Quantitative Aspects of the Analysis of the Monoclonal Antibody Trastuzumab Using High-Performance Liquid Chromatography Coupled with Electrospray Mass Spectrometry.* J Pharma Biomed Anal **46**, 449-455 (2008).

Appendix

Appendix A Proteins identified in lung microsomes

Table i: Identified proteins in cytochrome P450 molecular weight range of samples P1–P8

Accession number	Protein description	protein score	protein coverage	no of peptides	P1	P2	P3	P4	P5	P6	P7	P8
IPI00745872	ALB Isoform 1 of Serum albumin precursor	12221	59.3	31	x	x	x	x	x	x	x	x
IPI00003362	HSPA5 HSPA5 protein	5518	40.3	25	x	x	x	x	x	x	x	x
IPI00010796	P4HB Protein disulfe-isomerase precursor	2447	37.6	20	x	x	x	x	x	x	x	x
IPI00020984	CANX Calnexin precursor	1965	19.4	11	x				x	x	x	
IPI00003865	HSPA8 Isoform 1 of Heat shock cognate 71 kDa protein	1920	20.6	13	x	x	x	x	x	x	x	x
IPI00021440	PSPHL; ACTG1 Actin, cytoplasmic 2	1830	51.2	19			x	x				
IPI00021439	ACTB Actin, cytoplasmic 1	1816	51.2	13			x	x				
IPI00221226	ANXA6 Annexin A6	1622	34.5	21	x		x					
IPI00789477	LTF 73 kDa protein	1400	35.9	20				x	x	x		
IPI00219365	MSN Moesin	1393	28.9	16	x		x	x	x		x	x
IPI00301277	HSPA1L Heat shock 70 kDa protein 1L	1379	12	7	x			x		x	x	x
IPI00010180	CES1 Isoform 1 of Liver carboxylesterase 1 precursor	1369	29.8	16	x	x	x					x
IPI00027230	HSP90B1 Endoplasmin precursor	1270	24.4	20	x		x	x	x	x	x	
IPI00025874	RPN1 Dolichyl-diphosphooligosaccharide-protein glycosyltransferase 67 kDa subunit precursor	1254	20.7	13	x	x	x	x	x	x	x	
IPI00021428	ACTA1 Actin, alpha skeletal muscle	1187	33.4	15			x					
IPI00304925	HSPA1B; HSPA1A Heat shock 70 kDa protein 1	1145	15.8	9	x		x			x		x
IPI00384938	IGHG1 Hypothetical protein DKFZp686N02209	960	21.6	9		x	x	x				x
IPI00025252	PDIA3 Protein disulfe-isomerase A3 precursor	956	29.1	14	x	x	x	x			x	x
IPI00746388	VIL2 Villin 2	906	17	11	x	x	x	x	x		x	
IPI00026314	GSN Isoform 1 of Gelsolin precursor	854	17.1	9	x				x	x	x	
IPI00019502	MYH9 Myosin-9	832	9.5	15		x	x	x	x		x	
IPI00479186	PKM2 Isoform M2 of Pyruvate kinase isozymes M1/M2	828	38.4	15							x	x
IPI00017367	RDX Radixin	810	14.6	8						x		
IPI00386879	IGHA1 CDNA FLJ14473 fis, clone MAMMA1001080, highly similar to Homo sapiens SNC73 protein (SNC73) mRNA	778	18.6	6					x		x	x
IPI00339269	HSPA6 Heat shock 70 kDa protein 6	763	6.2	4	x		x			x		
IPI00007244	MPO Isoform H17 of Myeloperoxase precursor	718	15	11		x		x	x	x	x	

Accession number	Protein description	protein score	protein coverage	no of peptides	P1	P2	P3	P4	P5	P6	P7	P8
IPI00022774	VCP Transitional endoplasmic reticulum ATPase	681	14.5	9	x				x	x	x	
IPI00465248	ENO1 Isoform alpha-enolase of Alpha-enolase	673	33.4	13								x
IPI00019912	HSD17B4 Peroxisomal multifunctional enzyme type 2	664	19.7	12	x			x		x	x	
IPI00009904	PDIA4 Protein disulfe-isomerase A4 precursor	635	18.1	11	x		x	x		x	x	
IPI00020599	CALR Calreticulin precursor	604	21.1	7	x			x		x	x	x
IPI00553177	SERPINA1 Alpha-1-antitrypsin precursor	575	25.8	8								x
IPI00334775	HSP90AB1 85 kDa protein	529	6.4	12	x				x	x		x
IPI00024067	CLTC Isoform 1 of Clathrin heavy chain 1	527	7	11	x				x	x	x	
IPI00026154	PRKCSH Glucosidase 2 subunit beta precursor	524	13.4	8	x				x	x	x	
IPI00022463	TF Serotransferrin precursor	505	21.6	14	x				x	x	x	
IPI00550991	SERPINA3 Alpha-1-antichymotrypsin precursor	476	16.5	7	x	x		x		x	x	
IPI00168728	IGHM FLJ00385 protein (Fragment)	474	10.6	6		x		x				x
IPI00399007	IG1IG2 Hypothetical protein DKFZp686I04196 (Fragment)	472	17	8				x				x
IPI00465436	CAT Catalase	467	23.9	11	x				x	x	x	x
IPI00010471	LCP1 Plastin-2	440	22.6	13				x				x
IPI00100980	EHD2 EH domain-containing protein 2	435	12.9	8	x				x	x	x	
IPI00470467	POR NADPH-cytochrome P450 reductase	425	12.1	7	x					x	x	
IPI00000105	MVP Major vault protein	392	10.1	9	x	x	x	x	x	x	x	
IPI00008483	MAOA Amine oxase [flavin-containing] A	374	16.5	8	x	x		x	x	x	x	
IPI00013808	ACTN4 Alpha-actinin-4	366	5.9	5	x				x			
IPI00008494	ICAM1 Intercellular adhesion molecule 1 precursor	366	12.6	6	x				x		x	
IPI00383751	CALRETICULINCALCIUM binding protein (Fragment)	353	15.4	4	x		x		x		x	x
IPI00440493	ATP5A1 ATP synthase subunit alpha, mitochondrial precursor	336	17.2	9		x	x	x				
IPI00012303	SELENBP1 Selenium-binding protein 1	331	26.7	10								x
IPI00012728	ACSL1 Isoform 1 of Long-chain-fatty-acid-CoA ligase 1	323	9.7	7	x			x		x	x	
IPI00013683	TUBB3 Tubulin beta-3 chain	320	16.4	7			x					
IPI00007752	TUBB2C Tubulin beta-2C chain	318	13.3	6			x					x
IPI00218343	TUBA1C Tubulin alpha-1C chain	315	16.9	7		x	x					
IPI00217766	SCARB2 Lysosome membrane protein 2	310	6.9	4	x					x	x	
IPI00004457	AOC3 Membrane copper amine oxase	302	9.4	7	x				x			
IPI00743857	MYH11 smooth muscle myosin heavy chain 11 isoform SM1B	289	3.4	6		x			x			
IPI00022361	SLC4A1 Band 3 anion transport protein	281	8.1	7					x			
IPI00219077	LTA4H Isoform 1 of Leukotriene A-4 hydrolase	280	10.3	6							x	x
IPI00337335	MYH14 myosin, heavy chain 14 isoform 1	278	2.6	5					x			
IPI00299571	PDIA6 Isoform 2 of Protein disulfe-isomerase A6 precursor	271	10.8	5			x	x				x
IPI00218914	ALDH1A1 Retinal dehydrogenase 1	270	16.4	8								x
IPI00141318	CKAP4 Isoform 1 of Cytoskeleton-associated protein 4	264	11.1	6			x	x				

Accession number	Protein description	protein score	protein coverage	no of pepides	P1	P2	P3	P4	P5	P6	P7	P8
IPI00028031	ACADVL Isoform 1 of Very-long-chain specific acyl-CoA dehydrogenase, mitochondrial precursor	245	9.3	6	x						x	x
IPI00022462	TFRC Transferrin receptor protein 1	240	7.5	6					x	x	x	
IPI00303476	ATP5B ATP synthase subunit beta, mitochondrial precursor	236	15.3	8		x	x					x
IPI00477090	IGHM IGHM protein	233	9.6	7					x	x		
IPI00220834	XRCC5 ATP-dependent DNA helicase 2 subunit 2	226	4.2	4	x					x	x	
IPI00216171	ENO2 Gamma-enolase	216	7.1	3								x
IPI00009747	LSS Lanosterol synthase	210	7.5	5						x		
IPI00257508	DPYSL2 Dihydropyriminase-related protein 2	205	13.3	7								x
IPI00022143	FAM62A Isoform 1 of Protein FAM62A	199	1.2	1	x					x	x	
IPI00013475	TUBB2A Tubulin beta-2A chain	199	13.9	6		x						
IPI00328170	GCS1 Mannosyl-oligosaccharide glucosidase	197	4.1	4	x					x	x	
IPI00022200	COL6A3 alpha 3 type VI collagen isoform 1 precursor	195	1	4	x	x						
IPI00031522	HADHA Trifunctional enzyme subunit alpha, mitochondrial precursor	188	4.5	4	x		x			x	x	x
IPI00010418	MYO1C Myosin-Ic	176	4.6	5	x	x			x			
IPI00396485	EEF1A1 Elongation factor 1-alpha 1	175	10.4	6								x
IPI00028635	RPN2 Ribophorin II	172	7.3	5			x		x		x	
IPI00412259	LOC389901 similar to ATP-dependent DNA helicase 2 subunit 1 (ATP-dependent DNA helicase II 70 kDa subunit) (Lupus Ku autoantigen protein p70) (Ku70) (70 kDa subunit of Ku antigen) (Thyro-lupus autoantigen) (TLAA) (CTC box-b(	172	9.9	5						x		
IPI00641706	TUBB6 46 kDa protein	164	9	4		x	x					
IPI00030275	TRAP1 Heat shock protein 75 kDa, mitochondrial precursor	163	2	1	x				x			x
IPI00291792	ITGB2 Integrin beta-2 precursor	159	3.1	3	x				x	x		
IPI00022275	SACM1L SAC1 suppressor of actin mutations 1-like	159	9.4	5	x			x		x	x	
IPI00006663	ALDH2 Aldehyde dehydrogenase, mitochondrial precursor	157	14.7	8								x
IPI00030828	H6PD 90 kDa protein	152	4.4	3	x					x	x	
IPI00022895	A1BG Alpha-1B-glycoprotein precursor	149	9.3	5	x			x	x	x		
IPI00021338	DLAT Dihydrolipoyllysine-resue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial precursor	149	6.7	5						x	x	
IPI00177728	CNDP2 Cytosolic non-specific dipeptidase	147	7.4	3								x
IPI00397526	MYH10 Isoform 1 of Myosin-10	145	3	6					x			
IPI00165261	SCFD1 Sec1 family domain-containing protein 1	140	2.3	1	x					x		
IPI00022624	GPRC5A Retinoic ac-induced protein 3	136	3.6	1					x			
IPI00384280	PCYOX1 Prenylcysteine oxase precursor	135	9.9	4		x		x				
IPI00295542	NUCB1 Nucleobindin-1 precursor	134	7.4	3	x			x			x	

Accession number	Protein description	protein score	protein coverage	no of peptides	P1	P2	P3	P4	P5	P6	P7	P8
IPI00021891	FGG Isoform Gamma-B of Fibrinogen gamma chain precursor	129	7.1	3								x
IPI00328156	MAOB Amine oxase [flavin-containing] B	126	6.5	3	x							
IPI00294739	SAMHD1 SAM domain and HD domain-containing protein 1	124	6.4	3						x	x	x
IPI00156689	VAT1 Synaptic vesicle membrane protein VAT-1 homolog	124	6.4	2								x
IPI00032311	LBP Lipopolysaccharide-binding protein precursor	123	8.3	3	x					x		
IPI00294578	TGM2 Isoform 1 of Protein-glutamine gamma-glutamyltransferase 2	122	7.6	4						x	x	
IPI00017454	Tubulin alpha-4 chain	122	10.8	2		x		x				
IPI00643920	TKT Transketolase	121	5.6	2								x
IPI00220991	AP2B1 Hypothetical protein DKFZp781K0743	119	1.6	1	x					x		
IPI00029741	ITGB5 Integrin beta-5 precursor	116	1.1	1	x				x			
IPI00470649	NCLN Isoform 1 of Nicalin precursor	114	4.3	2	x						x	
IPI00027497	GPI Glucose-6-phosphate isomerase	106	8.4	3								x
IPI00298497	FGB Fibrinogen beta chain precursor	98	3.1	1								x
IPI00021048	FER1L3 Isoform 1 of Myoferlin	98	1.2	2							x	
IPI00014810	AGER Isoform 1 of Advanced glycosylation end product-specific receptor precursor	97	11.6	3	x		x	x				
IPI00302925	CCT8 Chaperonin containing TCP1, subunit 8 (Theta) variant	95	8.7	4						x		
IPI00004503	LAMP1 lysosomal-associated membrane protein 1	94	4.8	2						x	x	
IPI00292993	BPI Bacterical/permeability-increasing protein	94	4.9	2				x				
IPI00064193	TXNDC10 Isoform 1 of Protein disulfe-isomerase TXNDC10 precursor	93	3.5	1						x		
IPI00027509	MMP9 Matrix metalloproteinase-9 precursor	90	3.7	2					x			
IPI00472102	HSPD1 61 kDa protein	88	5.9	2								x
IPI00028064	CTSG Cathepsin G precursor	88	5.1	1						x		
IPI00215998	CD63 CD63 antigen	86	4.2	1	x						x	
IPI00217143	SDHA SDHA protein	85	2.7	1	x						x	
IPI00296157	RETSAT Isoform 1 of All-trans-retinol 13, 14-reductase precursor	85	4.1	2						x	x	
IPI00020987	PRELP Prolargin precursor	85	5.8	2	x					x		
IPI00031131	C20orf3 Adipocyte plasma membrane-associated protein	84	5.5	2				x				
IPI00479877	ALDH9A1 aldehyde dehydrogenase 9A1	82	1.9	1								x
IPI00386755	ERO1L ERO1-like protein alpha precursor	82	5.6	2	x						x	
IPI00011631	ZW10 Centromere/kinetochore protein zw10 homolog	81	1.7	1						x		
IPI00002406	BCAM Lutheran blood group glycoprotein precursor	80	1.8	1					x		x	
IPI00028091	ACTR3 Actin-like protein 3	79	5	2								x
IPI00000690	AIFM1 Isoform 1 of Apoptosis-inducing factor 1, mitochondrial precursor	79	2.1	1							x	

Accession number	Protein description	protein score	protein coverage	no of peptides	P1	P2	P3	P4	P5	P6	P7	P8
IPI00295400	WARS Tryptophanyl-tRNA synthetase, cytoplasmic	79	6.8	2								x
IPI00152418	CD55 Decay-accelerating factor splicing variant 4	78	2.3	1		x						
IPI00384122	DLST Full-length cDNA 5-PRIME end of clone CS0DB006YE12 of Neuroblastoma of Homo sapiens (Fragment)	74	11.7	3			x					
IPI00419916	ALPL Alkaline phosphatase, tissue-nonspecific isozyme precursor	74	2.1	1							x	
IPI00554541	ILVBL ilvB (bacterial acetolactate synthase)-like	74	7	2	x			x				
IPI00027547	DCD Dermcin precursor	71	20	2				x				
IPI00290553	ALDH1L1 10-formyltetrahydrofolate dehydrogenase	71	1.2	1								x
IPI00456492	CROCC Isoform 1 of Rootletin	70	0.4	4	x	x						
IPI00012889	hCG2025128;SFTPA1 Pulmonary surfactant-associated protein A1 precursor	70	5.2	1							x	
IPI00021302	SUSD2 Sushi domain-containing protein 2 precursor	69	4.7	2	x						x	
IPI00171626	AYTL2 1-acylglycerophosphocholine O-acyltransferase 1	65	2.2	1				x				
IPI00027412	CEACAM6 Carcinoembryonic antigen-related cell adhesion molecule 6 precursor	63	6.1	1	x					x		
IPI00009030	LAMP2 Isoform LAMP-2A of Lysosome-associated membrane glycoprotein 2 precursor	62	2	1						x	x	
IPI00555874	Actin-like protein (Fragment)	61	10.7	1			x					
IPI00022058	DDEF2 Isoform 1 of Development and differentiation-enhancing factor 2	61	0.8	1				x				x
IPI00027626	CCT6A T-complex protein 1 subunit zeta	61	2.3	1						x		
IPI00011454	GANAB Isoform 2 of Neutral alpha-glucosidase AB precursor	60	2	2	x							
IPI00418471	VIM Vimentin	60	2.1	1			x					x
IPI00013623	SLC27A3 Isoform 1 of Long-chain fatty acid transport protein 3	60	2.6	2	x					x		
IPI00217882	SORT1 Sortilin precursor	60	1.8	1	x							
IPI00024804	ATP2A1 Isoform SERCA1B of Sarcoplasmic/endoplasmic reticulum calcium ATPase 1	59	1.5	1					x			
IPI00646625	TAP1 transporter 1, ATP-binding cassette, sub-family B	59	3.5	2	x					x		
IPI00008967	TBXAS1 thromboxane A synthase 1 (platelet, cytochrome P450, family 5, subfamily A) isoform TXS-II	58	2.2	1	x	x						
IPI00001952	ENDOD1 Endonuclease domain-containing 1 protein precursor	58	2.6	1				x				
IPI00297084	DDOST dolichyl-diphosphooligosaccharide-protein glycosyltransferase precursor	57	4.8	2		x						
IPI00302592	FLNA filamin A, alpha	57	0.4	1					x			

Accession number	Protein description	protein score	protein coverage	no of peptides	P1	P2	P3	P4	P5	P6	P7	P8
IPI00021807	GBA Isoform Long of Glucosylceramase precursor	55	2.4	1							x	
IPI00171903	HNRPM Isoform 1 of Heterogeneous nuclear ribonucleoprotein M	54	2.2	1	x							
IPI00216008	G6PD Isoform Long of Glucose-6-phosphate 1-dehydrogenase	53	4.1	2								x
IPI00006070	CD34 Isoform CD34-F of Hematopoietic progenitor cell antigen CD34 precursor	53	5.2	1						x		
IPI00017184	EHD1 EH domain-containing protein 1	53	4.1	2						x		
IPI00022488	HPX Hemopexin precursor	53	4.5	1								x
IPI00550523	DKFZP564J0863 Isoform 1 of Atlastin-3	52	1.8	1				x				
IPI00032140	SERPINH1 Serpin H1 precursor	52	2.9	1		x						
IPI00073454	COL6A2 Isoform 2C2A~ of Collagen alpha-2(VI) chain precursor	51	1.3	1	x							
IPI00016801	GLUD1 Glutamate dehydrogenase 1, mitochondrial precursor	51	1.6	1								x
IPI00333619	ALDH3A2 Isoform 1 of Fatty aldehyde dehydrogenase	51	4.1	2		x	x					
IPI00441344	GLB1 Beta-galactosase precursor	51	2.1	1	x							
IPI00027493	SLC3A2;LOC442497 4F2 cell-surface antigen heavy chain	51	4.9	2						x	x	
IPI00012503	PSAP Isoform Sap-mu-0 of Proactivator polypeptide precursor	50	1.7	1							x	
IPI00382480	Ig heavy chain V-III region BRO	49	15.8	1								x
IPI00783024	Myosin-reactive immunoglobulin heavy chain variable region (Fragment)	49	14.5	1								x
IPI00021766	RTN4 Isoform 1 of Reticulon-4	48	1.1	1			x					
IPI00007461	DPY19L1 dpy-19-like 1	48	1.5	2							x	
IPI00375498	TTN titin isoform novex-1	48	0	1								x
IPI00216256	WDR1 Isoform 2 of WD repeat protein 1	48	3	1								x
IPI00328243	PLD3 Phospholipase D3	48	2	1								x
IPI00056505	NT5C3L 5~-nucleotase, cytosolic III-like	48	2.7	1				x				
IPI00301263	CAD protein	47	0.4	1								x
IPI00022256	AP2M1 AP-2 complex subunit mu-1	47	1.8	1			x					
IPI00550069	RNH1 Ribonuclease inhibitor	46	2.4	1								x
IPI00298828	APOH Beta-2-glycoprotein 1 precursor	45	2.6	1			x					
IPI00009896	EPHX1 Epoxe hydrolase 1	45	2.2	1								x
IPI00179330	UBB; RPS27A; UBC ubiquitin and ribosomal protein S27a precursor	45	10.3	1	x							
IPI00398958	LOC387867 similar to 40S ribosomal protein SA	45	5.1	1			x					
IPI00045223	HEJ1	44	15.6	1				x				
IPI00292858	ECGF1 Thymine phosphorylase precursor	43	3.9	1								x
IPI00163644	OSBPL8 Oxysterol-binding protein	42	1.2	1						x		
IPI00218718	ACSL6 acyl-CoA synthetase long-chain family member 6 isoform b	42	1.4	1							x	
IPI00419237	LAP3 Isoform 1 of Cytosol aminopeptidase	42	2.3	1								x

Accession number	Protein description	protein score	protein coverage	no of peptides	P1	P2	P3	P4	P5	P6	P7	P8
IPI00020042	PSMC4 Isoform 1 of 26S protease regulatory subunit 6B	42	2.9	1						x		
IPI00478292	SPTAN1 285 kDa protein	42	0.4	1					x			
IPI00414320	ANXA11 Annexin A11	42	1.8	1								x
IPI00794237	CALR Protein	41	10.1	1	x							
IPI00007765	HSPA9 Stress-70 protein, mitochondrial precursor	41	1.8	1						x		
IPI00465233	EIF3S6IP DJ1014D13.1 protein	41	2.1	1						x		
IPI00032826	ST13 Hsc70-interacting protein	41	2.7	1								x
IPI00018219	TGFBI Transforming growth factor-beta-induced protein ig-h3 precursor	41	1.8	1							x	
IPI00291930	CLINT1 Isoform 1 of Clathrin interactor 1	40	1.8	1						x		
IPI00455607	Uncharacterised 53 kDa protein	40	1.7	1						x		
IPI00298237	TPP1 Isoform 1 of Tripeptyl-peptidase 1 precursor	40	2.5	1								x
IPI00796116	ALDH3A1 42 kDa protein	39	3.2	1								x
IPI00247063	MME Neprilysin	39	1.2	1	x							
IPI00000875	EEF1G Elongation factor 1-gamma	39	3	1								x
IPI00043560	TSNARE1 T-SNARE domain containing 1	38	1.4	1			x					
IPI00028520	NDUFV1 Isoform 1 of NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial precursor	38	1.7	1			x					
IPI00030122	UTS2 Isoform 1 of Urotensin-2 precursor	38	6.5	1							x	
IPI00455056	LOC442710 similar to EF-hand calcium binding domain 1	37	10.7	1					x			
IPI00009960	IMMT Isoform 1 of Mitochondrial inner membrane protein	37	1.2	1						x		
IPI00011229	CTSD Cathepsin D precursor	37	2.2	1							x	
IPI00219525	PGD 6-phosphogluconate dehydrogenase, decarboxylating	37	2.1	1								x
IPI00168262	GLT25D1 CDNA PSEC0241 fis, clone NT2RP3000234, moderately similar to Homo sapiens cerebral cell adhesion molecule mRNA	36	1.4	1					x			
IPI00032291	C5 Complement C5 precursor	36	0.4	1	x							
IPI00743121	SMPD4 neutral sphingomyelinase 3 isoform 1	36	1	1	x							
IPI00478003	A2M Alpha-2-macroglobulin precursor	35	0.9	1					x			

Table ii: Comparison of proteins identified in whole gel lane of samples P3 and P4 using method 2.

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00021439	ACTB Actin, cytoplasmic 1	10342	430	59.5	19	1982	73	44.8	16
IPI00021440	PSPHL; ACTG1 Actin, cytoplasmic 2	10008	420	59.5	19	1856	68	44.8	16
IPI00023006	ACTC1 Actin, alpha cardiac muscle 1	6599	274	52.5	15	1439	55	40.8	12
IPI00455315	ANXA2 Annexin A2	4042	156	156	30	1682	56	51	24
IPI00019502	MYH9 Myosin-9	3616	156	156	70	280	12	4.8	7
IPI00654755	HBB Hemoglobin subunit beta	1777	76	76	11	442	19	82.3	10
IPI00743857	MYH11 smooth muscle myosin heavy chain 11 isoform SM1B	1530	65	65	28	81	3	1.6	3
IPI00329801	ANXA5 Annexin A5	1290	47	47	11	327	14	22.8	7
IPI00745872	ALB Isoform 1 of Serum albumin precursor	1157	55	55	22	948	37	21.8	14
IPI00012889	hCG_2025128;SFTPA1 Pulmonary surfactant-associated protein A1 precursor	1130	33	33	6	227	4	5.2	1
IPI00218918	ANXA1 Annexin A1	1112	32	32	11	218	8	13.3	3
IPI00397526	MYH10 Isoform 1 of Myosin-10	1030	53	53	28				
IPI00005171	HLA-DRA HLA class II histocompatibility antigen, DR alpha chain precursor	970	37	37	9	88	3	16.5	2
IPI00793199	ANXA4 annexin IV	948	25	25	14				
IPI00829896	HBD Hemoglobin Lepore-Baltimore (Fragment)	928	39	39	7	254	11	65.7	6
IPI00009236	CAV1 Isoform Alpha of Caveolin-1	911	42	42	7	268	12	28.7	5
IPI00784458	FBN1 312 kDa protein	906	42	42	30				
IPI00016342	RAB7A Ras-related protein Rab-7a	904	34	34	7				
IPI00453473	Histone H4	904	42	42	6	288	11	40.8	4
IPI00024933	RPL12 60S ribosomal protein L12	869	31	31	4	92	2	15.2	2
IPI00013698	ASAH1 Ac ceramase precursor	842	41	41	11				
IPI00328415	CYB5R3 Isoform 1 of NADH-cytochrome b5 reductase 3	842	33	33	9	181	8	14.6	4
IPI00646304	PPIB peptylprolyl isomerase B precursor	816	46	46	7	168	6	19	4
IPI00329598	HSD17B11 Dehydrogenase/reductase SDR family member 8 precursor	804	22	22	9				
IPI00303476	ATP5B ATP synthase subunit beta, mitochondrial precursor	799	29	29	13	104	3	6.8	3
IPI00215719	RPL18 60S ribosomal protein L18	782	25	25	5	355	9	25.5	4
IPI00008964	RAB1B Ras-related protein Rab-1B	779	30	30	6	179	6	15.9	3
IPI00218343	TUBA1C Tubulin alpha-1C chain	765	39	39	7				
IPI00410714	HBA2; HBA1 Hemoglobin subunit alpha	764	35	35	4	146	8	31.7	4

IPI human portein accession	Protein description	P3 method 2 protein score	P3 method 2 protein matches	P3 method 2 protein coverage	P3 method 2 no of peptides	P4 method 2 protein score	P4 method 2 protein matches	P4 method 2 protein coverage	P4 method 2 no of peptides
IPI00008530	RPLIPI0 60S acic ribosomal protein IPI0	741	28	28	6	132	5	13.6	4
IPI00180675	TUBA1A Tubulin alpha-1A chain	733	38	38	7				
IPI00009896	EPHX1 Epoxe hydrolase 1	716	41	41	18	121	7	11.4	5
IPI00397828	20 kDa protein	676	38	38	4	37	3	18.3	1
IPI00005719	RAB1A Isoform 1 of Ras-related protein Rab-1A	644	26	26	26				
IPI00027230	HSP90B1 Endoplasmin precursor	639	35	35	14	514	17	14.3	10
IPI00011654	TUBB Tubulin beta chain	636	36	36	9	350	14	15.5	6
IPI00299571	PDIA6 Isoform 2 of Protein disulfe-isomerase A6 precursor	614	21	21	7	188	3	5.7	2
IPI00025252	PDIA3 Protein disulfe-isomerase A3 precursor	565	24	24	13	471	19	29.1	13
IPI00335168	MYL6 Isoform Non-muscle of Myosin light polypeptide 6	563	27	27	9	155	6	27.8	4
IPI00337335	MYH14 myosin, heavy chain 14 isoform 1	555	24	24	9				
IPI00016339	RAB5C Ras-related protein Rab-5C	545	17	17	8	162	5	16.7	3
IPI00016513	RAB10 Ras-related protein Rab-10	539	21	21	3				
IPI00030179	RPL7 60S ribosomal protein L7	537	14	14	4	177	4	8.5	2
IPI00156689	VAT1 Synaptic vesicle membrane protein VAT-1 homolog	536	16	16	7	52	1	3.1	1
IPI00007752	TUBB2C Tubulin beta-2C chain	529	30	30	9				
IPI00004457	AOC3 Membrane copper amine oxidase	526	30	30	7	116	5	4.6	4
IPI00375676	FTL Ferritin light polypeptide variant	526	41	41	5				
IPI00465121	GNAI2 Galphai2 protein	526	20	20	7	147	4	10.1	3
IPI00796861	LOC653949; hCG_31916 WUGSC:H_RG054D04.1 protein	522	14	14	4				
IPI00291467	SLC25A6 ADP/ATP translocase 3	519	32	32	10	123	6	16.4	4
IPI00024282	RAB8B Ras-related protein Rab-8B	493	19	19	3	139	5	14.5	3
IPI00003348	GNB2 Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta 2	489	20	20	7	194	6	16.8	5
IPI00028481	RAB8A Ras-related protein Rab-8A	489	18	18	3				
IPI00217775	CD74 Isoform Short of HLA class II histocompatibility antigen gamma chain	485	17	17	5				
IPI00020418	RRAS Ras-related protein R-Ras precursor	485	12	12	3				
IPI00003362	HSPA5 HSPA5 protein	481	26	26	15	666	22	31.8	17
IPI00024095	ANXA3 Annexin A3	481	20	20	12	264	10	31.3	8
IPI00291928	RAB14 Ras-related protein Rab-14	479	18	18	4	249	9	27.4	5
IPI00024067	CLTC Isoform 1 of Clathrin heavy chain 1	462	26	26	11	245	10	6.3	9
IPI00033494	MRLC2 Myosin regulatory light chain	459	21	21	5	116	4	18	3
IPI00219682	STOM Erythrocyte band 7 integral membrane protein	453	9	9	3	235	9	30.6	6
IPI00026268	GNB1 Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta 1	453	17	17	7	185	5	13.2	4
IPI00032140	SERPINH1 Serpin H1 precursor	451	15	15	11				

IPI human portein accession	Protein description	P3 method 2 protein score	P3 method 2 protein matches	P3 method 2 protein coverage	P3 method 2 no of peptides	P4 method 2 protein score	P4 method 2 protein matches	P4 method 2 protein coverage	P4 method 2 no of peptides
IPI00013683	TUBB3 Tubulin beta-3 chain	451	25	25	5				
IPI00300096	RAB35 Ras-related protein Rab-35	450	17	17	2	121	4	10.9	2
IPI00394882	RAB15 Isoform 1 of Ras-related protein Rab-15	450	17	17	2	121	4	10.4	
IPI00011229	CTSD Cathepsin D precursor	446	24	24	8	40	2	4.1	1
IPI00013485	RPS2 40S ribosomal protein S2	445	14	14	3	86	4	10.2	3
IPI00297084	DDOST dolichyl-diphosphooligosaccharide-protein glycosyltransferase precursor	443	14	14	7	111	5	9.2	3
IPI00418471	VIM Vimentin	438	20	20	10	86	5	10.3	5
IPI00013808	ACTN4 Alpha-actinin-4	434	20	20	9	155	2	3.1	2
IPI00178083	TPM3 29 kDa protein	432	20	20	10				
IPI00007188	SLC25A5 ADP/ATP translocase 2	431	28	28	10	100	6	15.8	3
IPI00028055	TMED10 Transmembrane emp24 domain-containing protein 10 precursor	423	14	14	4	168	6	13.7	3
IPI00419424	IGKV1-5 IGKV1-5 protein	419	13	13	4	39	1	5.8	1
IPI00026314	GSN Isoform 1 of Gelsolin precursor	418	17	17	11	135	4	4.9	3
IPI00000105	MVP Major vault protein	418	22	22	14	473	20	19.7	16
IPI00025512	HSPB1 Heat-shock protein beta-1	413	17	17	5	68	2	13.2	1
IPI00027252	PHB2 Prohibitin-2	412	18	18	9	45	1	3.7	1
IPI00009634	SQRDL Sulfide:quinone oxoreductase, mitochondrial precursor	410	12	12	9				
IPI00020984	CANX Calnexin precursor	410	25	25	10	251	9	13.2	7
IPI00023526	RAB6A Isoform 1 of Ras-related protein Rab-6A	402	14	14	3				
IPI00021812	AHNAK AHNAK nucleoprotein isoform 1	402	16	16	7				
IPI00022891	SLC25A4 ADP/ATP translocase 1	393	24	24	9	98	4	13.4	3
IPI00007676	HSD17B12 Estradiol 17-beta-dehydrogenase 12	391	18	16	7	136	3	10.9	3
IPI00746734	HLA class II histocompatibility antigen, DR-1 beta chain precursor	390	16	17	6				
IPI00555744	RPL14 RPL14 protein	390	17	9	3	117	4	13.2	2
IPI00477489	MIA;RAB4B Isoform 1 of Ras-related protein Rab-4B	385	9	14	3				
IPI00448725	RAB4B RAB4B protein	384	14	15	3				
IPI00014055	NAPSA Napsin-A precursor	374	15	20	3				
IPI00221226	ANXA6 Annexin A6	374	20	21	13				
IPI00299573	RPL7A 60S ribosomal protein L7a	373	21	13	4				
IPI00005614	SPTBN1 Isoform Long of Spectrin beta chain, brain 1	371	13	12	9	50	2	0.8	1
IPI00554521	FTH1 Ferritin heavy chain	371	12	19	4				
IPI00010779	TPM4 Isoform 1 of Tropomyosin alpha-4 chain	368	19	13	6				
IPI00025874	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase 67 kDa	368	13	17	7	105	4	5.4	3

IPI human portein accession	Protein description	P3 method 2 protein score	P3 method 2 protein matches	P3 method 2 protein coverage	P3 method 2 no of peptides	P4 method 2 protein score	P4 method 2 protein matches	P4 method 2 protein coverage	P4 method 2 no of peptides
	subunit precursor								
IPI00006865	SEC22B Vesicle-traflicking protein SEC22b	367	17	12	4	64	1	6.5	1
IPI00219219	LGALS1 Galectin-1	365	12	21	8	112	4	23.7	3
IPI00020436	RAB11B Ras-related protein Rab-11B	360	21	17	8	84	4	13.8	3
IPI00304925	HSPA1B; HSPA1A Heat shock 70 kDa protein 1	360	17	15	7				
IPI00015148	RAP1B Ras-related protein Rap-1b precursor	358	15	11	6	74	4	11.4	2
IPI00641950	GNB2L1 Lung cancer oncogene 7	357	11	11	7				
IPI00022300	METTL7A Methyltransferase-like protein 7A precursor	355	11	9	3	145	4	12.3	2
IPI00010796	P4HB Protein disulfide-isomerase precursor	351	9	24	10	575	15	23.4	10
IPI00219018	GAPDH Glyceraldehyde-3-phosphate dehydrogenase	342	24	10	6	54	2	4.5	1
IPI00217030	RPS4X 40S ribosomal protein S4, X isoform	340	10	13	5	122	5	6.1	1
IPI00024145	VDAC2 Isoform 1 of Voltage-dependent anion-selective channel protein 2	339	13	12	7	92	2	7.2	2
IPI00013508	ACTN1 Alpha-actinin-1	339	12	16	7				
IPI00003865	HSPA8 Isoform 1 of Heat shock cognate 71 kDa protein	335	16	14	6	227	8	13.6	6
IPI00413108	RPSA Ribosomal protein SA	333	14	9	5				
IPI00216308	VDAC1 Voltage-dependent anion-selective channel protein 1	333	9	11	36				
IPI00031169	RAB2A Ras-related protein Rab-2A	332	11	15	7	207	4	18.9	3
IPI00182933	CYB5A Isoform 2 of Cytochrome b5	330	15	15	5	62	3	31.6	2
IPI00219365	MSN Moesin	319	15	24	12	181	12	14.7	7
IPI00219153	RPL22 60S ribosomal protein L22	310	24	11	2				
IPI00220278	MYL9 Myosin regulatory light chain 2, smooth muscle isoform	298	11	15	6				
IPI00304612	RPL13A 60S ribosomal protein L13a	297	15	11	5	108	3	14.8	3
IPI00010418	MYO1C Myosin-Ic	289	11	14	8				
IPI00298237	TPP1 Isoform 1 of Tripeptyl-peptidase 1 precursor	287	14	11	3				
IPI00792510	SFTPC 18 kDa protein	284	11	12	2	68	2	18	2
IPI00419942	TPSB2 TPSB2 protein	281	12	12	6				
IPI00007067	C9orf19 Golgi-associated plant pathogenesis-related protein 1	280	12	6	2				
IPI00376798	RPL11 Isoform 1 of 60S ribosomal protein L11	278	6	9	2	146	4	12.9	2
IPI00298289	RTN4 Isoform 2 of Reticulon-4	277	9	10	3				
IPI00017334	PHB Prohibitin	275	10	10	7	64	1	4.4	1
IPI00000230	TPM1 tropomyosin 1 alpha chain isoform 2	265	10	10	3				
IPI00183968	TPM3 tropomyosin 3 isoform 1	265	10	11	4				
IPI00009253	NAPA Alpha-soluble NSF attachment protein	263	11	9	9	95	2	8.5	2

IPI human portein accession	Protein description	P3 method 2 protein score	P3 method 2 protein matches	P3 method 2 protein coverage	P3 method 2 no of peptides	P4 method 2 protein score	P4 method 2 protein matches	P4 method 2 protein coverage	P4 method 2 no of peptides
IPI00100980	EHD2 EH domain-containing protein 2	262	9	9	7				
IPI00028064	CTSG Cathepsin G precursor	259	9	12	3	145	6	12.9	3
IPI00177817	ATP2A2 Isoform SERCA2A of Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	256	12	9	5				
IPI00003925	Isoform 1 of Pyruvate dehydrogenase E1 component subunit beta, mitochondrial precursor	256	9	7	5				
IPI00302592	FLNA filamin A, alpha	256	7	9	7	65	1	0.5	1
IPI00216587	RPS8 40S ribosomal protein S8	256	9	9	5	123	3	13.5	2
IPI00218820	TPM2 Isoform 3 of Tropomyosin beta chain	256	9	9	2				
IPI00010274	TPSAB1 Isoform 1 of Tryptase alpha-1 precursor	255	9	10	5	60	1	5.5	1
IPI00017510	MT-CO2 Cytochrome c oxidase subunit 2	254	10	11	3	65	2	4.4	1
IPI00305383	Ubiquinol-cytochrome-c reductase complex core protein 2, mitochondrial precursor	253	11	8	5				
IPI00440493	ATP5A1 ATP synthase subunit alpha, mitochondrial precursor	253	8	14	11	53	1	2.9	1
IPI00005202	PGRMC2 Membrane-associated progesterone receptor component 2	251	14	10	4				
IPI00339269	HSPA6 Heat shock 70 kDa protein 6	250	10	10	3				
IPI00011253	RPS3 40S ribosomal protein S3	250	10	14	6	186	6	19.8	4
IPI00377005	RcTPM3 (Fragment)	249	14	12	8				
IPI00021805	MGST1 Microsomal glutathione S-transferase 1	248	12	12	3				
IPI00176903	PTRF Isoform 1 of Polymerase I and transcript release factor	246	12	10	4				
IPI00016786	CDC42 Isoform 2 of Cell division control protein 42 homolog precursor	243	10	9	1				
IPI00031804	VDAC3 Isoform 1 of Voltage-dependent anion-selective channel protein 3	240	9	8	3				
IPI00018534	HIST1H2BL Histone H2B type 1-L	240	8	14	4	109	7	34.9	3
IPI00738107	LOC730415 HLA class II histocompatibility antigen, DRB1-1 beta chain precursor	240	14	9	4				
IPI00746388	VIL2 Villin 2	239	9	17	7				
IPI00465248	ENO1 Isoform alpha-enolase of Alpha-enolase	239	17	7	7				
IPI00301021	SSR1 Isoform 1 of Translocon-associated protein subunit alpha precursor	239	7	10	2				
IPI00221091	RPS15A 40S ribosomal protein S15a	237	10	12	2	95	3	13.8	2
IPI00301280	TMEM43 Transmembrane protein 43	230	12	10	6				
IPI00009407	DAD1 Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit DAD1	229	10	8	2				
IPI00220578	GNAI3 Guanine nucleotide-binding protein G	226	8	12	4				

IPI human portein accession	Protein description	P3 method 2 protein score	P3 method 2 protein matches	P3 method 2 protein coverage	P3 method 2 no of peptides	P4 method 2 protein score	P4 method 2 protein matches	P4 method 2 protein coverage	P4 method 2 no of peptides
IPI00847347	LOC730415 MHC class II antigen	224	12	9	3				
IPI00376964	COL6A3 alpha 3 type VI collagen isoform 2 precursor	220	9	13	10				
IPI00006482	Isoform Long of Sodium/potassium-transporting ATPase subunit alpha-1 precursor	216	13	9	6	96	2	2.3	2
IPI00550165	DHRS7B dehydrogenase/reductase (SDR family) member 7B	215	9	4	1				
IPI00339384	RDH11 Isoform 1 of Retinol dehydrogenase 11	215	4	6	4	38	1	3.8	1
IPI00299084	TMEM33 Transmembrane protein 33	214	6	7	4				
IPI00017344	RAB5B Ras-related protein Rab-5B	213	7	9	5				
IPI00026569	HLA class I histocompatibility antigen, A-1 alpha chain precursor	211	9	5	4	76	2	7.4	2
IPI00472855	HLA class I histocompatibility antigen, A-30 alpha chain precursor	211	5	5	4				
IPI00031131	C20orf3 Adipocyte plasma membrane-associated protein	211	5	12	9	120	6	11.8	4
IPI00012011	CFL1 Cofilin-1	209	12	7	3	107	2	13.9	2
IPI00027193	CLIC5 Isoform 2 of Chloride intracellular channel protein 5	208	7	7	5				
IPI00412579	RPL10A 60S ribosomal protein L10a	205	7	7	5				
IPI00152409	AGR3 Anterior gradient protein 3 homolog precursor	204	7	14	3	54	2	13.3	1
IPI00007426	ARL6IP5 PRA1 family protein 3	203	14	9	3	72	2	4.8	1
IPI00023542	TMED9 transmembrane emp24 protein transport domain containing 9	202	9	9	3	75	1	4.7	1
IPI00412855	hCG_2040224 similar to ribosomal protein L18a	200	9	9	3	115	3	15.3	2
IPI00026202	RPL18A 60S ribosomal protein L18a	200	9	9	3	115	3	15.3	2
IPI00004672	HLA-A 41 kDa protein	199	9	5	4				
IPI00472510	HLA class II histocompatibility antigen, DQ(1) alpha chain precursor	194	5	12	2				
IPI00019912	HSD17B4 Peroxisomal multifunctional enzyme type 2	193	12	6	6	63	1	2.2	1
IPI00009904	PDIA4 Protein disulfe-isomerase A4 precursor	192	6	9	6	38	1	1.6	1
IPI00021841	APOA1 Apolipoprotein A-I precursor	192	9	5	4	45	1	4.9	1
IPI00815838	RNASE2 MHC class II antigen (Fragment)	190	5	8	3				
IPI00217563	ITGB1 Isoform Beta-1A of Integrin beta-1 precursor	189	8	8	7	170	6	5	4
IPI00024911	ERP29 Endoplasmic reticulum protein ERp29 precursor	188	8	7	3	59	1	4.6	1
IPI00031691	RPL9 60S ribosomal protein L9	187	7	7	3	64	2	5.2	1
IPI00384938	IGHG1 Hypothetical protein DKFZp686N02209	181	7	8	5	86	5	8.9	4
IPI00219729	SLC25A11 Mitochondrial 2-oxoglutarate/malate carrier protein	180	8	7	6				
IPI00003918	RPL4 60S ribosomal protein L4	179	7	9	3	88	4	6.1	2
IPI00170692	VAPA Vesicle-associated membrane protein-associated protein A	179	9	6	2				

IPI human portein accession	Protein description	P3 method 2 protein score	P3 method 2 protein matches	P3 method 2 protein coverage	P3 method 2 no of peptides	P4 method 2 protein score	P4 method 2 protein matches	P4 method 2 protein coverage	P4 method 2 no of peptides
IPI00654689	HLA-A HLA class I histocompatibility antigen,	176	6	5	5				
IPI00010397	LOC730415 HLA class II DR-beta (HLA-DR B) precursor	176	5	7	3				
IPI00215893	HMOX1 Heme oxygenase 1	172	7	5	4				
IPI00023510	RAB5A Ras-related protein Rab-5A	171	5	9	5				
IPI00300299	SPCS3 Signal peptidase complex subunit 3	171	9	3	1				
IPI00477090	IGHM IGHM protein	170	3	4	4				
IPI00221089	RPS13 40S ribosomal protein S13	170	4	12	2	45	2	11.3	2
IPI00000816	YWHAE 14-3-3 protein epsilon	169	12	12	4	50	3	8.6	2
IPI00219757	GSTP1 Glutathione S-transferase P	169	12	4	3	39	1	4.8	1
IPI00003935	HIST2H2BE Histone H2B type 2-E	168	4	10	4				
IPI00002412	PPT1 Palmitoyl-protein thioesterase 1 precursor	166	10	5	2				
IPI00215884	SFRS1 Isoform ASF-1 of Splicing factor, arginine/serine-rich 1	165	5	8	5	73	2	7.7	2
IPI00004657	HLA-A HLA class I histocompatibility antigen, B-7 alpha chain precursor	164	8	5	5	57	2	7.5	1
IPI00215997	CD9 antigen	164	5	8	3	40	2	4.4	1
IPI00027438	FLOT1 Flotillin-1	160	8	4	4				
IPI00026154	PRKCSH Glucosidase 2 subunit beta precursor	159	4	8	7	66	2	3.2	2
IPI00215914	ARF1 ADP-ribosylation factor 1	158	8	7	4	101	3	11.6	2
IPI00218200	BCAP31 B-cell receptor-associated protein 31	156	7	5	3	45	2	5.1	2
IPI00013847	UQCRC1 Ubiquinol-cytochrome-c reductase complex core protein 1, mitochondrial precursor	156	5	6	4				
IPI00008433	RPS5 40S ribosomal protein S5	154	6	2	1	36	1	8.3	1
IPI00022975	ALOX5AP Arachonate 5-lipoxygenase-activating protein	153	2	10	3	42	1	5	1
IPI00333619	ALDH3A2 Isoform 1 of Fatty aldehyde dehydrogenase	153	10	6	5	91	2	5.2	2
IPI00021263	YWHAZ 14-3-3 protein zeta/delta	150	6	11	4	75	3	9	2
IPI00010180	CES1 Isoform 1 of Liver carboxylesterase 1 precursor	150	11	8	5				
IPI00183695	S100A10 Protein S100-A10	149	8	9	2				
IPI00452747	LOC653566 hypothetical protein LOC653566	148	9	4	2	52	1	4.9	1
IPI00215918	ARF4 ADP-ribosylation factor 4	148	4	5	2				
IPI00329332	STX12 Synin-12	146	5	6	5	41	1	5.4	1
IPI00002230	AADACL1 arylacetamide deacetylase-like 1	145	6	4	4				
IPI00016670	C11orf59 UPF0404 protein C11orf59	143	4	2	1	47	1	8.1	1
IPI00018871	ARL8B ADP-ribosylation factor-like protein 8B	143	2	8	5				
IPI00103082	HLA-DPB1 HLA class II histocompatibility antigen, DP(W4) beta chain precursor	139	8	7	3	61	1	4.3	1
IPI00550021	RPL3 60S ribosomal protein L3	137	7	6	3				

		P3 method 2				P1 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00473031	ADH1B Alcohol dehydrogenase 1B	135	6	6	5	159	3	9.3	3
IPI00022624	GPRC5A Retinoic ac-induced protein 3	135	6	3	1	251	5	3.6	1
IPI00478292	SPTAN1 285 kDa protein	135	3	9	7	94	4	0.8	2
IPI00005186	HLA-DQB1 HLA class II histocompatibility antigen, DQ(1) beta chain precursor	133	9	6	2				
IPI00465279	LOC730415 HLA class II histocompatibility antigen, DRB1-7 beta chain precursor	133	6	4	3	67	1	3.8	1
IPI00031697	TMEM109 Transmembrane protein 109 precursor	133	4	3	1	95	3	4.9	4
IPI00000190	CD81 CD81 antigen	133	3	4	2				
IPI00025796	NDUFS3 NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial precursor	132	4	3	3				
IPI00465431	LGALS3 Galectin-3	132	3	7	5				
IPI00013955	MUC1 Isoform 1 of Mucin-1 precursor	132	7	6	2				
IPI00022463	TF Serotransferrin precursor	131	6	4	4	66	2	3.6	2
IPI00420108	DLST Dihydrolipoyllysine-resue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial precursor	131	4	2	2				
IPI00022793	HADHB Trifunctional enzyme beta subunit, mitochondrial precursor	130	2	6	6				
IPI00061977	IGHA1 IGHA1 protein	130	6	3	3	84	3	8.2	2
IPI00257059	TMEM173 Transmembrane protein 173	130	3	6	2	54	1	2.4	1
IPI00339223	FN1 Isoform 3 of Fibronectin precursor	129	6	3	2				
IPI00470467	POR NADPH-cytochrome P450 reductase	126	3	7	3				
IPI00029046	KIAA0152 Uncharacterized protein KIAA0152 precursor	126	7	4	3				
IPI00218474	ENO3 Beta-enolase	125	4	2	2				
IPI00739274	LOC652614 similar to HLA class I histocompatibility antigen, A-11 alpha chain precursor	125	2	3	3				
IPI00011937	PRDX4 Peroxiredoxin-4	124	3	6	3				
IPI00026942	ERLIN2 Isoform 1 of Erlin-2 precursor	124	6	3	3	123	2	6.5	2
IPI00441414	GANAB Isoform 3 of Neutral alpha-glucosidase AB precursor	122	3	8	8				
IPI00220739	PGRMC1 Membrane-associated progesterone receptor component 1	122	8	5	4	35	1	4.1	
IPI00013917	RPS12 40S ribosomal protein S12	122	5	7	2				
IPI00000874	PRDX1 Peroxiredoxin-1	122	7	7	3	61	2	10.6	2
IPI00798360	CIP29 18 kDa protein	122	7	7	1	53	2	4.3	1
IPI00008483	MAOA Amine oxidase [flavin-containing] A	121	7	7	7	75	3	7.6	3
IPI00022391	APCS Serum amyloid P-component precursor	121	7	5	5				
IPI00654777	EIF3S5 Eukaryotic translation initiation factor 3 subunit 5	119	5	2	2				

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00022018	DPM1 Dolichol-phosphate mannosyltransferase	117	2	2	1				
IPI00025049	M6PR Cation-dependent mannose-6-phosphate receptor precursor	117	2	2	1				
IPI00005737	SURF4 Isoform 1 of Surfeit locus protein 4	116	2	2	1				
IPI00296190	C10orf58 Uncharacterized protein C10orf58 precursor	116	2	4	1				
IPI00029731	RPL35A 60S ribosomal protein L35a	115	4	8	4	36	1	8.2	1
IPI00296083	SFTPB Pulmonary surfactant-associated protein B precursor	114	8	7	3				
IPI00010896	CLIC1 Chloride intracellular channel protein 1	112	7	3	3				
IPI00549343	VAMP3 Vesicle-associated membrane protein 3	111	3	4	2	92	2	17	1
IPI00024670	REEP5 Receptor expression-enhancing protein 5	111	4	5	2				
IPI00026994	PRAF2 PRA1 family protein 2	111	5	4	1				
IPI00217519	RALA Ras-related protein Ral-A precursor	109	4	7	3	45	1	6.8	1
IPI00019038	LYZ Lysozyme C precursor	109	7	4	2	117	3	8.1	1
IPI00306419	MBOAT5 Membrane-bound O-acyltransferase domain-containing protein 5	109	4	4	3				
IPI00007910	SLC34A2 Isoform 1 of Sodium-dependent phosphate transport protein 2B	108	4	3	2	46	1	1.7	1
IPI00029133	ATP5F1 ATP synthase B chain, mitochondrial precursor	108	3	4	4				
IPI00004503	LAMP1 lysosomal-associated membrane protein 1	107	4	4	2	92	3	4.8	2
IPI00028635	RPN2 Ribophorin II	107	4	5	3				
IPI00021923	FAM3C Protein FAM3C precursor	105	5	2	2				
IPI00217766	SCARB2 Lysosome membrane protein 2	104	2	4	2	69	1	2.1	1
IPI00008529	RPLP2 60S acidic ribosomal protein P2	103	2	3	1	44	1	10.4	1
IPI00291136	COL6A1 Collagen alpha-1 (VI) chain precursor	102	3	5	3	299	9	7.4	6
IPI00009507	SYPL1 Isoform 1 of Synaptophysin-like protein 1	102	5	2	1	59	2	4.2	1
IPI00022202	SLC25A3 Isoform A of Phosphate carrier protein, mitochondrial precursor	101	2	4	2	48	1	3.3	1
IPI00010270	RAC2 Ras-related C3 botulinum toxin substrate 2 precursor	101	4	4	3	57	2	11.5	2
IPI00294472	TMED5 Transmembrane emp24 domain-containing protein 5 precursor	101	4	3	2				
IPI00014810	AGER Isoform 1 of Advanced glycosylation end product-specific receptor precursor	99	3	6	4	96	2	3.7	1
IPI00141318	CKAP4 Isoform 1 of Cytoskeleton-associated protein 4	99	6	3	2	80	4	7.8	4
IPI00029625	FLOT2 Flotillin-2	99	3	2	2				
IPI00027466	CA4 Carbonic anhydrase 4 precursor	99	2	4	4	48	1	4.8	1
IPI00022895	A1BG Alpha-1B-glycoprotein precursor	98	4	4	3	52	1	2.2	1

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00478410	ATP5C1 Isoform Liver of ATP synthase gamma chain, mitochondrial precursor	97	4	4	4				
IPI00022395	C9 Complement component C9 precursor	97	4	2	2				
IPI00329593	ADPGK Isoform 2 of ADP-dependent glucokinase	96	2	1	1				
IPI00007611	ATP5O ATP synthase O subunit, mitochondrial precursor	96	1	4	4	51	1	5.2	1
IPI00219446	PEBP1 Phosphatylethanolamine-binding protein 1	96	4	2	1				
IPI00009950	LMAN2 Vesicular integral-membrane protein VIP36 precursor	96	2	6	4				
IPI00465028	TPI1 Triosephosphate isomerase 1 variant	95	6	3	3				
IPI00012728	ACSL1 Isoform 1 of Long-chain-fatty-acid-CoA ligase 1	94	3	6	6	40	1	1.3	1
IPI00815988	HLA-DRB4; HLA-DRB1; HLA-DRB5; HLA-DRB3; LOC730415 HLA-DRB1 protein (Fragment)	94	6	5	2				
IPI00007755	RAB21 Ras-related protein Rab-21	94	5	1	1				
IPI00220487	ATP5H Isoform 1 of ATP synthase D chain, mitochondrial	94	1	6	4				
IPI00183603	DC2 DC2 protein	94	6	3	1				
IPI00020599	CALR Calreticulin precursor	93	3	3	3	56	2	5	2
IPI00154742	IGL@ IGL@ protein	93	3	2	2	63	2	4.3	1
IPI00063130	UNQ501 MBC3205	92	2	2	1				
IPI00026182	CAPZA2 F-actin capping protein subunit alpha-2	92	2	3	3				
IPI00000494	RPL5 60S ribosomal protein L5	92	3	3	3				
IPI00022774	VCP Transitional endoplasmic reticulum ATPase	91	3	3	3	112	4	6.5	4
IPI00005087	TMOD3 Tropomodulin-3	91	3	2	2				
IPI00299719	TCIRG1 Isoform Long of Vacuolar proton translocating ATPase 116 kDa subunit a isoform 3	91	2	4	2				
IPI00006957	DHRS7 Isoform 1 of Dehydrogenase/reductase SDR family member 7 precursor	90	4	4	4	73	2	5.6	2
IPI00007797	LOC728641; FABP5; LOC731043 Fatty acid-binding protein, epermal	90	4	4	3				
IPI00219155	RPL27 60S ribosomal protein L27	89	4	4	2				
IPI00025491	EIF4A1 Eukaryotic initiation factor 4A-I	88	4	2	2				
IPI00002460	ANXA7 Isoform 1 of Annexin A7	88	2	1	1				
IPI00012772	RPL8 60S ribosomal protein L8	88	1	3	2	78	2	10.5	1
IPI00002149	SAR1B GTP-binding protein SAR1b	87	3	3	1				
IPI00005969	CAPZA1 F-actin capping protein subunit alpha-1	87	3	3	3	45	1	3.5	1
IPI00019385	SSR4 Translocon-associated protein delta subunit precursor	86	3	3	2				
IPI00465436	CAT Catalase	86	3	5	5	44	1	2.5	1

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00419585	PPIAL3; PPIA; LOC654188 Peptyl-prolyl cis-trans isomerase A	86	5	6	3	98	4	10.9	2
IPI00026570	COX7A2 Cytochrome *c* oxidase polypeptide VIIa-liver/heart, mitochondrial precursor	85	6	2	1				
IPI00289876	STX7 Isoform 1 of Synin-7	85	2	4	2				
IPI00221384	COL12A1 Isoform 2 of Collagen alpha-1 (XII) chain precursor	84	4	2	1				
IPI00221232	GNG12 Guanine nucleotide-binding protein G(I)/G(S)/G(O) gamma-12 subunit precursor	84	2	3	1				
IPI00221088	RPS9 40S ribosomal protein S9	84	3	6	4	48	2	4.6	1
IPI00022275	SACM1L SAC1 suppressor of actin mutations 1-like	83	6	4	4	35	1	1.5	1
IPI00027848	MRC1 Macrophage mannose receptor 1 precursor	83	4	3	3	39	1	0.8	1
IPI00020557	LRP1 Low-density lipoprotein receptor-related protein 1 precursor	83	3	3	2				
IPI00014898	PLEC1 Isoform 1 of Plectin-1	83	3	1	1				
IPI00736170	25 kDa protein	82	1	3	2				
IPI00017454	Tubulin alpha-4 chain	82	3	3	2				
IPI00470674	CYB5R1 NADH-cytochrome b5 reductase 1	82	3	2	2	58	1	3.6	1
IPI00003635	ERGIC1 Isoform 2 of Endoplasmic reticulum-Golgi intermediate compartment protein 1	81	2	3	2				
IPI00014587	CLTA Isoform Brain of Clathrin light chain A	81	3	3	3				
IPI00005809	SDPR Serum deprivation-response protein	81	3	4	2				
IPI00099595	HSD17B6 hydroxysteroid (17-beta) dehydrogenase 6	80	4	2	2				
IPI00219291	ATP5J2 Isoform 2 of ATP synthase f chain, mitochondrial	79	2	2	1				
IPI00030919	MAP2K1IP1 Mitogen-activated protein kinase kinase 1-interacting protein 1	79	2	3	1				
IPI00401264	TXNDC4 Thioredoxin domain-containing protein 4 precursor	78	3	2	2	81	3	7.6	3
IPI00056414	MAL2 Protein MAL2	78	2	3	1				
IPI00014958	PON2 Isoform 1 of Serum paraoxonase/arylesterase 2	77	3	1	1	48	1	4.5	1
IPI00016621	AP2A2 AP-2 complex subunit alpha-2	75	1	2	2				
IPI00171421	C8orf55 Uncharacterized protein C8orf55 precursor	75	2	4	2				
IPI00005721	DEFA1; LOC728358;LOC653600 Neutrophil defensin 1 precursor	74	4	3	2				
IPI00215983	CA1 Carbonic anhydrase 1	74	3	2	2				
IPI00018364	RAP2B Ras-related protein Rap-2b precursor	74	2	1	1				
IPI00025285	FLJ25715;ATP6V1G1 Vacuolar ATP synthase subunit G 1	73	1	1	1				
IPI00032150	CDS2 Isoform 1 of Phosphatate cytidylyltransferase 2	73	1	2	1				

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00171626	AYTL2 1-acylglycerophosphocholine O-acyltransferase 1	73	2	3	2	65	1	2.2	1
IPI00014149	TTC35 Tetratricopeptide repeat protein 35	72	3	2	1				
IPI00030362	PLP2 Proteolipid protein 2	72	2	2	1	51	1	8.6	1
IPI00010438	SNAP23 Isoform SNAP-23a of Synaptosomal-associated protein 23	72	2	1	1				
IPI00009030	LAMP2 Isoform LAMP-2A of Lysosome-associated membrane glycoprotein 2 precursor	72	1	2	1	71	4	2	1
IPI00003856	ATP6V1E1 Vacuolar ATP synthase subunit E 1	72	2	2	2				
IPI00291792	ITGB2 Integrin beta-2 precursor	72	2	4	2	148	6	4.6	2
IPI00004454	DPM3 Isoform 1 of Dolichol-phosphate mannosyltransferase subunit 3	72	4	2	1				
IPI00073454	COL6A2 Isoform 2C2A~ of Collagen alpha-2(VI) chain precursor	71	2	5	2				
IPI00221178	TPD52L2 Isoform 2 of Tumor protein D54	71	5	1	1				
IPI00021785	COX5B Cytochrome c oxidase subunit 5B, mitochondrial precursor	71	1	2	2				
IPI00010402	SH3BGRL3 Hypothetical protein	71	2	1	1				
IPI00002535	FKBP2 FK506-binding protein 2 precursor	70	1	3	1				
IPI00291328	NDUFV2 NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial precursor	70	3	2	1				
IPI00002406	BCAM Lutheran blood group glycoprotein precursor	70	2	2	1				
IPI00000643	BAG2 BAG family molecular chaperone regulator 2	69	2	1	1				
IPI00034159	ATP6V0D1 Vacuolar ATP synthase subunit d 1	69	1	2	2	68	1	3.7	1
IPI00104128	SEC11A Signal peptidase complex catalytic subunit SEC11A	69	2	4	1	54	2	9.5	1
IPI00026105	SCP2 Isoform SCPx of Nonspecific lip-transfer protein	69	4	2	2				
IPI00009747	LSS Lanosterol synthase	69	2	2	2				
IPI00329352	NOMO1 Nodal modulator 1 precursor	69	2	2	2				
IPI00556589	LOC650788 similar to 40S ribosomal protein S28	68	2	3	1				
IPI00306301	PDHA1 Pyruvate dehydrogenase E1 component alpha subunit, somatic form, mitochondrial precursor	68	3	2	2				
IPI00009235	SSR3 Translocon-associated protein subunit gamma	68	2	1	1				
IPI00025447	EEF1A1 EEF1A1 protein	68	1	2	2				
IPI00018415	TM9SF2 Transmembrane 9 superfamily protein member 2 precursor	66	2	2	1				
IPI00001952	ENDOD1 Endonuclease domain-containing 1 protein precursor	66	2	1	1	58	1	2.6	1
IPI00102685	MYADM Myelo-associated differentiation marker	65	1	3	1				

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00026964	UQCRFS1 Ubiquinol-cytochrome c reductase iron-sulfur subunit, mitochondrial precursor	64	3	3	2				
IPI00031522	HADHA Trifunctional enzyme subunit alpha, mitochondrial precursor	64	3	4	2				
IPI00155168	PTPRC Protein tyrosine phosphatase, receptor type, C	64	4	2	2	50	1	1	1
IPI00009829	CPA3 Mast cell carboxypeptidase A precursor	64	2	2	2				
IPI00221372	AGPAT2 Isoform 2 of 1-acyl-sn-glycerol-3-phosphate acyltransferase beta	64	2	1	1				
IPI00384280	PCYOX1 Prenylcysteine oxase precursor	64	1	2	2	95	4	9.5	3
IPI00216691	PFN1 Profilin-1	63	2	3	2	101	2	18.6	1
IPI00001754	F11R Junctional adhesion molecule A precursor	63	3	3	2				
IPI00219217	LDHB L-lactate dehydrogenase B chain	63	3	2	1				
IPI00289819	IGF2R Cation-independent mannose-6-phosphate receptor precursor	62	2	2	2				
IPI00828094	ABCA8 protein	62	2	2	1				
IPI00015972	COX6C Cytochrome *c* oxidase polypeptide VIc precursor	62	2	2	2				
IPI00294159	SLC25A1 Tricarboxylate transport protein, mitochondrial precursor	62	2	1	1				
IPI00328170	GCS1 Mannosyl-oligosaccharide glucosidase	61	1	1	1				
IPI00418495	CD36 Platelet glycoprotein 4	61	1	1	1				
IPI00006579	COX4I1 Cytochrome c oxidase subunit 4 isoform 1, mitochondrial precursor	61	1	2	1				
IPI00216298	TXN Thioredoxin	61	2	1	1				
IPI00182533	RPL28 60S ribosomal protein L28	60	1	2	1				
IPI00011302	CD59 CD59 glycoprotein precursor	59	2	1	1				
IPI00376206	HSD17B13 Isoform 2 of 17-beta hydroxysteroid dehydrogenase 13 precursor	59	1	1	1				
IPI00029741	ITGB5 Integrin beta-5 precursor	59	1	3	1	107	4	1.1	1
IPI00026530	LMAN1 ERGIC-53 protein precursor	59	3	2	2				
IPI00303568	PTGES2 Prostaglandin E synthase 2	57	2	1	1				
IPI00293590	MGLL monoglyceride lipase isoform 1	57	1	1	1				
IPI00006072	SEC61G Protein transport protein SEC61 subunit gamma	56	1	1	1				
IPI00027463	S100A6 Protein S100-A6	56	1	2	1				
IPI00645518	CDIPT Isoform 1 of CDP-diacylglycerol-inositol 3-phosphatyltransferase	56	2	1	1				
IPI00291878	SFTPD Pulmonary surfactant-associated protein D precursor	54	1	1	1	60	1	4	1
IPI00305551	GNA11 Guanine nucleotide-binding protein subunit alpha-11	54	1	2	2	54	2	7.2	2
IPI00216237	RPL36 60S ribosomal protein L36	54	2	1	1				

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00010471	LCP1 Plastin-2	54	1	1	1	51	1	2.1	1
IPI00026646	FCGRT IgG receptor FcRn large subunit p51 precursor (Fragment)	53	1	3	1				
IPI00016703	DHCR24 24-dehydrocholesterol reductase precursor	53	3	3	1				
IPI00294911	SDHB Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial precursor	53	3	1	1				
IPI00022361	SLC4A1 Band 3 anion transport protein	52	1	2	1				
IPI00165949	ARTS-1 type 1 tumor necrosis factor receptor shedding aminopeptidase regulator isoform a	52	2	1	1				
IPI00395887	TXNDC1 Thioredoxin domain-containing protein 1 precursor	52	1	2	1				
IPI00010845	NDUFS8 NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial precursor	52	2	2	1				
IPI00016381	RAB27A Isoform Long of Ras-related protein Rab-27A	52	2	2	2	87	4	19	4
IPI00012503	PSAP Isoform Sap-mu-0 of Proactivator polypeptide precursor	52	2	3	2				
IPI00220301	PRDX6 Peroxiredoxin-6	51	3	1	1	44	1	4	1
IPI00337494	SLC25A24 solute carrier family 25 member 24 isoform 1	51	1	2	2				
IPI00030397	BNIP1 Isoform 3 of Vesicle transport protein SEC20	50	2	1	1				
IPI00003588	EEF1E1 Eukaryotic translation elongation factor 1 epsilon-1	50	1	1	1				
IPI00294739	SAMHD1 SAM domain and HD domain-containing protein 1	50	1	1	1				
IPI00020472	TMEM111 Isoform 1 of Transmembrane protein 111	50	1	2	1				
IPI00443909	TMEM4 Isoform 1 of MIR-interacting saposin-like protein precursor	49	2	1	1				
IPI00006980	C14orf166 Protein C14orf166	49	1	2	2				
IPI00014577	RAB18 Ras-related protein Rab-18	49	2	1	1	35	1	5.3	1
IPI00218848	ATP5I ATP synthase, H+ transporting, mitochondrial F0 complex, subunit E	49	1	1	1				
IPI00056357	C19orf10 Uncharacterized protein C19orf10 precursor	48	1	2	1				
IPI00026824	HMOX2 Heme oxygenase 2	47	2	1	1				
IPI00100656	GPSN2 Isoform 1 of Synaptic glycoprotein SC2	46	1	2	2				
IPI00026328	TXNDC12 Thioredoxin domain-containing protein 12 precursor	46	2	2	1				
IPI00816799	Rheumatoid factor D5 light chain (Fragment)	45	2	1	1				
IPI00019407	NSDHL Sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating	45	1	2	1				
IPI00031091	EFHD1 EF-hand domain-containing protein 1	44	2	2	1				
IPI00027993	RAB25 Ras-related protein Rab-25	44	2	2	2				
IPI00022462	TFRC Transferrin receptor protein 1	44	2	3	3				

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00021828	CSTB Cystatin-B	44	3	1	1				
IPI00396321	LRRC59 Leucine-rich repeat-containing protein 59	44	1	1	1				
IPI00296909	PARP4 Poly [ADP-ribose] polymerase 4	44	1	1	1				
IPI00009342	IQGAP1 Ras GTPase-activating-like protein IQGAP1	43	1	1	1				
IPI00169383	PGK1 Phosphoglycerate kinase 1	42	1	1	1				
IPI00026241	BST2 Bone marrow stromal antigen 2 precursor	41	1	1	1				
IPI00550069	RNH1 Ribonuclease inhibitor	41	1	1	1				
IPI00008711	WFS1 Wolframin	40	1	1	1				
IPI00171903	HNRPM Isoform 1 of Heterogeneous nuclear ribonucleoprotein M	40	1	1	1				
IPI00305304	LASS2 LAG1 longevity assurance homolog 2	40	1	1	1				
IPI00003968	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial precursor	40	1	2	2				
IPI00299086	SDCBP Syntenin-1	40	2	1	1				
IPI00024920	ATP5D ATP synthase delta chain, mitochondrial precursor	39	1	1	1				
IPI00218850	SCAMP2 Secretory carrier-associated membrane protein 2	39	1	1	1				
IPI00018381	TLL1 Isoform 1 of Tolloid-like protein 1 precursor	39	1	1	1				
IPI00012555	FCN1 Ficolin-1 precursor	39	1	1	1				
IPI00007327	TAPBP Isoform 1 of Tapasin precursor	39	1	1	1	48	1	2.2	1
IPI00007730	C14orf1 Probable ergosterol biosynthetic protein 28	39	1	1	1				
IPI00166865	ZCD2 Zinc finger, CDGSH-type domain 2	39	1	1	1				
IPI00219673	GSTK1 Glutathione S-transferase kappa 1	39	1	1	1	46	1	7.1	1
IPI00296157	RETSAT Isoform 1 of All-trans-retinol 13, 14-reductase precursor	39	1	1	1				
IPI00386755	ERO1L ERO1-like protein alpha precursor	39	1	1	1				
IPI00292380	PIK3R3 Isoform 1 of Phosphoinosite 3-kinase regulatory	39	1	1	1				
IPI00329600	SCCPDH Probable saccharopine dehydrogenase	38	1	1	1				
IPI00414836	OSTF1 Osteoclast-stimulating factor 1	38	1	1	1				
IPI00002506	ALG5; LOC442122; C17orf13 Dolichyl-phosphate beta-glucosyltransferase	38	1	1	1				
IPI00005161	ARPC2 Actin-related protein 2/3 complex subunit 2	38	1	1	1				
IPI00059292	FLJ10292 Protein mago nashi homolog 2	38	1	1	1				
IPI00000690	AIFM1 Isoform 1 of Apoptosis-inducing factor 1, mitochondrial precursor	38	1	1	1				
IPI00412545	NDUFA5 Hypothetical protein DKFZp781K1356	37	1	1	1				
IPI00027500	RHOA Transforming protein RhoA precursor	37	1	1	1				
IPI00220099	STX3 Isoform B of Syntin-3	37	1	1	1				

IPI human portein accession	Protein description	P3 method 2 protein score	P3 method 2 protein matches	P3 method 2 protein coverage	P3 method 2 no of peptides	P4 method 2 protein score	P4 method 2 protein matches	P4 method 2 protein coverage	P4 method 2 no of peptides
IPI00219436	SEC11C Signal peptidase complex catalytic subunit SEC11C	37	1	1	1				
IPI00171412	SUMF2 Isoform 1 of Sulfatase-modifying factor 2 precursor	37	1	1	1				
IPI00022371	HRG Histine-rich glycoprotein precursor	37	1	1	1				
IPI00299000	PA2G4 Proliferation-associated protein 2G4	36	1	1	1				
IPI00011284	COMT Isoform Membrane-bound of Catechol O-methyltransferase	36	1	1	1	39	1	3	1
IPI00029623	PSMA6 Proteasome subunit alpha type 6	36	1	1	1				
IPI00013678	MTX1 mein 1 isoform 1	36	1	1	1				
IPI00019755	GSTO1 Glutathione transferase omega-1	36	1	1	1				
IPI00383000	26 kDa protein	36	1	1	1				
IPI00168262	GLT25D1 CDNA PSEC0241 fis, clone NT2RP3000234, moderately similar to Homo sapiens cerebral cell adhesion molecule mRNA	36	1	2	2				
IPI00152240	TMEM167 Transmembrane protein 167 precursor	35	2	1	1				
IPI00009976	TMED1 Transmembrane emp24 domain-containing protein 1 precursor	35	1	1	1				
IPI00029264	CYC1 Cytochrome c1 heme protein, mitochondrial precursor	35	1	1	1				
IPI00100247	TXNDC13 Thioredoxin domain-containing protein 13 precursor	34	1	1	1				
IPI00479743	POTE2 protein expressed in prostate, ovary, testis, and placenta 2					1163	33	7.3	5
IPI00789477	LTF 73 kDa protein					663	25	31.4	17
IPI00788271	LOC728320 similar to lactotransferrin					545	21	34.2	14
IPI00007244	MPO Isoform H17 of Myeloperoxase precursor					359	18	15.3	10
IPI00022200	COL6A3 alpha 3 type VI collagen isoform 1 precursor					260	11	2.8	8
IPI00829992	MYO1C myosin IC isoform b					260	11	12.5	9
IPI00027547	DCD Dermcidin precursor					260	11	22.7	3
IPI00473011	HBD; HBB Hemoglobin subunit delta					250	11	46.3	6
IPI00555692	ANXA4 protein					248	8	24.4	6
IPI00002459	ANXA6 annexin VI isoform 2					227	6	6.4	4
IPI00550991	SERPINA3 Alpha-1-antichymotrypsin precursor					188	6	10.9	4
IPI00166768	TUBA1C TUBA1C protein					182	9	14	4
IPI00555610	AHNAK Neuroblast differentiation-associated protein AHNAK					152	5	2.2	3
IPI00334775	HSP90AB1 85 kDa protein					140	4	3.4	2
IPI00075248	CALM3; CALM2; CALM1 Calmodulin					134	3	8.7	1
IPI00023504	RAB3A Ras-related protein Rab-3A					131	5	13.6	3
IPI00032808	RAB3D Ras-related protein Rab-3D					131	5	13.7	3

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00061114	RAB3C Ras-related protein Rab-3C					131	5	13.2	3
IPI00448725	RAB4B protein					128	4	19.8	2
IPI00301277	HSPA1L Heat shock 70 kDa protein 1L					124	4	8.1	4
IPI00012512	RRAS2 Ras-related protein R-Ras2 precursor					115	2	7.8	1
IPI00017601	CP Ceruloplasmin precursor					109	3	4	2
IPI00008494	ICAM1 Intercellular adhesion molecule 1 precursor					109	5	6.2	4
IPI00399007	IGHG2 Hypothetical protein DKFZp686I04196 (Fragment)					108	6	10.3	4
IPI00168728	IGHM FLJ00385 protein (Fragment)					104	5	8.4	1
IPI00001618	RAB39 Ras-related protein Rab-39A					101	3	5.1	1
IPI00021475	RAB33B Ras-related protein Rab-33B					101	3	4.8	1
IPI00472416	LA-A HLA class I histocompatibility antigen, B-45 alpha chain precursor					96	2	6.9	2
IPI00011134	HSPA7 Heat shock 70 kDa protein 7 (Fragment)					94	2	9.7	2
IPI00292993	BPI Bacterical/permeability-increasing protein					94	2	4.9	2
IPI00104074	CD163 Isoform 1 of Scavenger receptor cysteine-rich type 1 protein M130 precursor					89	3	2.1	2
IPI00299573	RPL7A 60S ribosomal protein L7a					87	3	7.9	2
IPI00304840	COL6A2 Isoform 2C2 of Collagen alpha-2(VI) chain precursor					84	3	3.2	3
IPI00022246	AZU1 Azurocidin precursor					84	4	8	2
IPI00847294	HLA-B; MICA; LOC730410; HLA-C; HLA-A29.1; HLA-A MHC class I antigen (Fragment)					82	2	14.4	2
IPI00018278	H2AFV Histone H2AV					81	3	7	1
IPI00290928	GNA13 Guanine nucleotide-binding protein alpha-13 subunit					79	2	2.9	1
IPI00017454	Tubulin alpha-4 chain					76	2	5	1
IPI00398135	hCG_21078 hypothetical protein LOC389435					75	1	7.4	1
IPI00299547	LCN2 Neutrophil gelatinase-associated lipocalin precursor					75	1	7.6	1
IPI00478003	A2M Alpha-2-macroglobulin precursor					70	2	1.5	2
IPI00171438	TXNDC5; MUTED Thioredoxin domain-containing protein 5 precursor					66	1	2.5	1
IPI00009650	LCN1 Lipocalin-1 precursor					64	2	12.5	2
IPI00027444	SERPINB1 Leukocyte elastase inhibitor					64	1	2.9	1
IPI00021302	SUSD2 Sushi domain-containing protein 2 precursor					63	1	1.5	1
IPI00045223	HEJ1					63	14	15.6	1
IPI00029015	FVT1 3-ketodihydrosphingosine reductase precursor					58	1	3.3	1
IPI00218251	TGM2 Isoform 2 of Protein-glutamine gamma-glutamyltransferase 2					58	1	1.5	1
IPI00514635	CLIC3 Protein					57	1	8.8	1

IPI human portein accession	Protein description	P3 method 2 protein score	P3 method 2 protein matches	P3 method 2 protein coverage	P3 method 2 no of peptides	P4 method 2 protein score	P4 method 2 protein matches	P4 method 2 protein coverage	P4 method 2 no of peptides
IPI00296259	TMED4 Transmembrane emp24 domain-containing protein 4 precursor					57	1	4.8	1
IPI00217966	LDHA Isoform 1 of L-lactate dehydrogenase A chain					57	1	3.3	1
IPI00013946	SYNGR2 Synaptogyrin-2					56	2	4.5	1
IPI00385264	Ig mu heavy chain disease protein					56	1	2.8	1
IPI00298497	FGB Fibrinogen beta chain precursor					55	1	3.1	1
IPI00000877	HYOU1 Hypoxia up-regulated protein 1 precursor					54	1	0.9	1
IPI00220834	XRCC5 ATP-dependent DNA helicase 2 subunit 2					54	1	1.6	1
IPI00026302	RPL31 60S ribosomal protein L31					53	1	7.2	1
IPI00179330	UBB; RPS27A; UBC ubiquitin and ribosomal protein S27a precursor					53	1	10.3	1
IPI00329389	RPL6 60S ribosomal protein L6					53	4	9.4	2
IPI00171611	HIST2H3A; HIST2H3C Histone H3.2					53	2	5.1	1
IPI00550523	DKFZP564J0863 Isoform 1 of Atlastin-3					52	1	1.8	1
IPI00025329	RPL19 60S ribosomal protein L19					52	1	4.6	1
IPI00477859	DMBT1 62 kDa protein					51	1	2.4	1
IPI00027509	MMP9 Matrix metalloproteinase-9 precursor					51	2	3.3	2
IPI00339224	FN1 Isoform 4 of Fibronectin precursor					50	2	1.4	1
IPI00018953	DPP4 Dipeptyl peptidase 4					50	1	1.8	1
IPI00021416	KATNAL2 40 kDa protein					49	1	3.3	1
IPI00220644	PKM2 Isoform M1 of Pyruvate kinase isozymes M1/M2					48	2	3.8	2
IPI00022058	DDEF2 Isoform 1 of Development and differentiation-enhancing factor 2					48	5	0.8	1
IPI00247583	RPL21; LOC731567; LOC729402 60S ribosomal protein L21					47	1	9.4	1
IPI00295542	NUCB1 Nucleobindin-1 precursor					45	1	2.2	1
IPI00012303	SELENBP1 Selenium-binding protein 1					45	1	2.1	1
IPI00305457	SERPINA1 PRO2275					44	1	15	1
IPI00397611	LOC388344 similar to ribosomal protein L13 isoform 1					43	2	5.2	1
IPI00217987	ITGAM Integrin alpha-M precursor					43	1	1.1	1
IPI00221092	RPS16 40S ribosomal protein S16					43	1	6.8	1
IPI00646498	GNAL 8 kDa protein					42	1	15.9	1
IPI00554541	ILVBL ilvB (bacterial acetolactate synthase)-like					42	1	3.3	1
IPI00217802	C12orf25 Uncharacterized protein C12orf25					42	1	0.9	1
IPI00021107	XYLB xylulokinase homolog					41	1	2.4	1
IPI00014424	EEF1A2 Elongation factor 1-alpha 2					40	1	1.7	1
IPI00218728	PAFAH1B1 Isoform 1 of Platelet-activating factor acetylhydrolase IB subunit alpha					39	1	1.7	1

		P3 method 2				P4 method 2			
IPI human portein accession	Protein description	protein score	protein matches	protein coverage	no of peptides	protein score	protein matches	protein coverage	no of peptides
IPI00157687	PECAM1 Isoform Delta15 of Platelet endothelial cell adhesion molecule precursor					39	1	1.8	1
IPI00017800	ABCA3 ATP-binding cassette sub-family A member 3					39	1	0.5	1
IPI00171230	ERC1 Isoform 2 of ELKS/RAB6-interacting/CAST family member 1					39	1	0.8	1
IPI00056505	NT5C3L 5~-nucleotidase, cytosolic III-like					39	4	2.7	1
IPI00257508	DPYSL2 Dihydropyrimidinase-related protein 2					39	1	1.9	1
IPI00027462	S100A9 Protein S100-A9					38	1	11.4	1
IPI00382470	HSP90AA1 Heat shock protein HSP 90-alpha 2					38	1	1.5	1
IPI00787363	SEC14L4; LOC731874; LOC730005 similar to SEC14p-like protein TAP3					37	1	2.3	1
IPI00437751	ACE Angiotensin-converting enzyme, somatic isoform precursor					37	1	0.8	1
IPI00016334	MCAM Isoform 1 of Cell surface glycoprotein MUC18 precursor					37	1	1.9	1
IPI00246827	C17orf28 89 kDa protein					36	2	1	1
IPI00004524	GCA Grancalcin					36	1	5.1	1
IPI00007756	RAB22A Ras-related protein Rab-22A					36	1	5.7	1
IPI00218646	CYBB Cytochrome b-245 heavy chain					36	1	1.8	1
IPI00215780	RPS19 40S ribosomal protein S19					36	1	6.2	1

Appendix B Overview of liver samples used for CYP1A2 and CYP2E1 quantification

Table iii: Overview of liver samples used for CYP1A2 and CYP2E1 quantification.

Sample	Tissue	Gender	Age	Tissue mass [g]	Microsomal protein concentration [µg/µl]	Total microsomal volume [ml]	Microsomal concentration in tissue [mg/g]	CYP2E1 [pmol/mg microsomal protein]	CYP1A2 [pmol/mg microsomal protein]
H1N	Liver control	M	41	0.63	6.09	2	12.18	129.76±2.81	
H1T	Primary liver tumour			0.66	3.4	0.3	1.55		
H2N	Liver control	F	72	0.5	8.85	0.4	7.08	144.85±41.63	
H2T	Primary liver tumour			0.5	3.2	0.25	1.6		
H3N	Liver control	F	79	0.64	4.72	0.75	5.53	200.51±51.99	
H3T	Colorectal metastases in liver			0.72	3.08	0.5	2.14		
H4N	Liver control	F	60	0.7	3.98	0.75	4.26	88.12±22.65	
H4T	Colorectal metastases in liver			0.63	2.4	0.5	1.9		
H6	Liver control	F	64	0.65	5.45	0.75	6.29		165.43±66.75
H7	Liver control	M	78	0.65	7.92	0.75	9.14		240.76±25.53
H8	Liver control	m	45	0.69	4.91	0.75	5.34		262.6±76.18

Appendix C Publications

Publications resulting from the work in this thesis

Seibert C., Wang Y., Davidson B. R., Fuller B. J., Patterson L. H., Griffiths, W. J.; *Quantification of CYP2E1 by stable isotope dilution mass spectrometry; In preparation*

Wang Y., Al-Gazzar A., Seibert C., Sharif A., Lane C. Griffiths W. J.; *Proteomic analysis of cytochromes P450: A mass spectrometry approach;* Biochemical Society Transactions. 2006 34(Pt 6):1246-51.

Conference contributions resulting from the work in the thesis

Wang Y., Davidson B. R., Fuller B. J., Griffiths W. J., Seibert C. ; *Quantification of cytochrome P450 isoforms in human liver tissue using stable isotope labelled standards;* Annual meeting of the British Mass Spectrometry Society, Edinburgh, September 2007

Seibert C., Wang Y., Al-Gazzar A., Davidson B. R., Fuller B. J., Griffiths W. J.; *Absolute quantification of cytochrome P450 enzymes in human liver tissue and liver tumour tissue;* International Mass spectrometry Conference, Prague, August 2006

Gazzar A., Seibert C., Griffiths W. J., Fuller B., Davidson B.; *A proteomic approach to the identification of pancreatic specific proteins;* Annual Meeting of the Society of Academic and Research Surgery, Edinburgh, January 2006

Lane C.S., Karu K., Seibert C., Griffiths W. J. and Patterson L.; *Relatative quantification of cytochrome P450 proteins in immune-deficient mice grafted with human tumors using ^{18}O stable isotope labelling;* Seventh International Symposium in Mass Spectrometry in the Health and Life Sciences, San Francisco, California, August 2005

Seibert C., Patterson L. H., Griffiths W. J.; *Identification of human lung microsomal proteins: A proteomic approach;* ULLA Summer School, Uppsala, Sweden July 2005

Seibert C., Brady S. M., Al-Gazzar A., Wang Y., Necal A., Patterson L.H., Turton J. A. Griffiths W. J.; *Identification of drug metabolising enzymes and disease markers: A proteomic approach;* Annual meeting of the British Mass Spectrometry Society, Derby, September 2004

Appendix D Accompanying CD

The accompanying CD contains data referred to in the text. The data is available as Excel- and .html file. A directory can be found in the file *index.html.*

Chapter 3:

LungP1toP8merge_CYParea	Identified proteins in CYP mass range area
2E1_experiment	Different amounts of recombinant CYP2E1 in-gel digests analysed on QTOF
Liver_experiment	Identified proteins in human liver microsomes derived from in-gel digests analysed on QTOF
massranges_comparison	Lung P3, band16 analysed using method 1 and 2: comparison of identified proteins
P3andP4method2	Identified proteins of in-gel digest of whole lane analysed of P3 and P4 using method 2
P3mehtod1	Identified proteins of in-gel digest of whole lane analysed of P3 using method 1

Chapter 4:

ControlliverH1NtoH4N	Proteins identified in Liver control samples H1N to H4N
TumourliverH1TtoH4T	Proteins identified in Liver tumour samples H1T to H4T

Chapter 5:

LiverH6toH8	Proteins identified in Liver control samples H6 to H8